CERTIFICATE OF VERIFICATION

I, Masayuki HARABAYASHI,
of c/o Morita & Associates, 5th Fl., Itabashi-Chuo Bldg.,
67-8, Itabashi 2-chome, Itabashi-ku, Tokyo 173-0004, Japan,
state that the attached document is a true and complete translation to the best of my knowledge of Japanese Patent Application No. 2001-156088.

Dated this 12th day of February, 2007

Signature of translator: M. Harabayashi

[DOCUMENT NAME] Patent Application
[REFERENCE NUMBER] YLS01001P
[SPECIAL REMARKS] Patent application stating that applicant seeks application of Section 30(1)of the Patent Law
[ADDRESSEE] Director-General of Patent Office
[INTERNATIONAL PATENT CLASSIFICATION] C07K 14/00
[INVENTOR]
[Address or Residence] 4-4-8, Yakumo, Meguro-ku, Tokyo
[Name] Shigeo OHNO
[APPLICANT]
[Address or Residence] 4-4-8, Yakumo, Meguro-ku, Tokyo
[Name] Shigeo OHNO
[ATTORNEY]
[Identification No.] 100090251
[Patent Attorney]
[Name] Kenichi MORITA
[INDICATION OF FEE]
[Deposit Account Number] 017813
[Fee] 21000
[LIST OF SUBMITTED DOCUMENTS]
[Name of Document] Specification 1
[Name of Document] Drawings 1
[Name of Document] Abstract 1
[REQUEST FOR A PROOF] Yes

[DOCUMENT NAME] Specification

[TITLE OF THE INVENTION] NOVEL SMG-1

[CLAIMS]

[Claim 1] (1) A polypeptide comprising an amino acid sequence consisting of 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, or (2) a polypeptide exhibiting an SMG-1 activity and comprising an amino acid sequence in which one or plural amino acids are deleted, substituted, and/or inserted at one or plural positions in an amino acid sequence consisting of 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2.

[Claim 2] A polynucleotide encoding the polypeptide according to claim 1.

[Claim 3] An expression vector comprising the polynucleotide according to claim 2.

[Claim 4] A cell transfected with the expression vector according to claim 3.

[Claim 5] An antibody which binds to the polypeptide according to claim 1.

[Claim 6] A method for screening a substance which inhibits an SMG-1 activity of the polypeptide according to claim 1, comprising the steps of:
bringing into contact (1) the polypeptide, (2) Upf1/SMG-2, and (3) a substance to be tested; and
carrying out phosphorylation under the conditions that the polypeptide is brought into contact with Upf1/SMG-2 and the test substance, and analyzing whether or not Upf1/SMG-2 is phosphorylated.

[DETAILED DESCRIPTION OF THE INVENTION]

[0001]

[Technical Field to which the Invention Pertains]

The present invention relates to SMG-1.

[0002]

[Prior Art]

In eukaryotes, although a promoter site is the same as that of a normal gene, a nonsense mutation mRNA, in which a codon in the inherent translational region of a gene is changed to a stop codon, is recognized and specifically degraded. One such mechanism for specific degradation is

nonsense mediated mRNA decay (NMD). As the genes relating to this mechanism, three genes (UPF1, UPF2, and UPF3) have been reported from yeast and seven genes (SMG-1 to SMG-7) from Caenorhabditis elegans. In mutant organisms of these genes, it has also been reported that the specific degradation of nonsense mutation mRNA is suppressed. In this connection, yeast UPF1 protein and C. elegans SMG-2 protein have a high homology between their amino acid sequences. Further, as a human gene and mouse gene having a high homology of the base sequence with the yeast UPF1 gene, Rent1/HUPF1 (hereinafter referred to simply as "human UPF1") has been isolated. It is shown that this gene complements the functions of UPF-1 in UPF-1 mutant yeast. Further, when expressing a mutant human UPF1 protein wherein the 844th arginine is mutated to cysteine in animal cells, a suppression of the specific degradation of nonsense mutated mRNA is seen. In this connection the mutants of these genes are not lethal, and are not believed to be genes required for survival.

[0003]

The UPF1/SMG-2 protein has a Zn finger motif and RNA helicase-like structure and is believed to function as a unit of the complex for degradation of mRNA. Further, other genes are believed to regulate, for example, the activity or location of this enzyme. In C. elegans, it has been reported that the SMG-2 protein is phosphorylated, and that in C. elegans of mutants of the genes of SMG-1, SMG-3, or SMG-4, the SMG-2 protein is not phosphorylated. Further, the base sequence of the cDNA of C. elegans SMG-1 has been reported. The SMG-1 protein has a kinase domain having a high homology with the kinase domain conserved as the family of the group of serine/threonine kinases known as phosphatidyl inositol kinase related kinases (PIKK) and is considered to be PIKK family. Further, a sequence believed to be fruit-fly SMG-1 has been reported from the base sequence of the fruit-fly genome gene. However, the base sequence of the SMG-1 gene of mammals, including humans, and the amino acid sequence of the SMG-1 protein encoding the same have not been elucidated.

[0004]

[Problems to be Solved by the Invention]

The present inventor engaged in intensive search with the object of obtaining a novel phosphatidyl inositol kinase (PIK) related kinase (PIKK) and, as a result, obtained a novel human SMG-1 protein and DNA encoding the same. Further, the present inventor showed that the human SMG-1 has an autophosphorylation activity and an activity of phosphorylating UPF1/SMG-2, and further immunoprecipitates together with UPF1/SMG-2, UPF2, and UPF3. From these facts, the present inventor proved that the human SMG-1 is a member of the surveillance complex which triggers the NMD, and that SMG-1 is actually essential for NMD in mammalian cells using point mutations of SMG-1. Further, the present inventor newly discovered that NMD can be suppressed by inhibiting human SMG-1. The present invention is based on these findings.

Therefore, the object of the present invention is to provide a novel phosphatidyl inositol kinase (PIK) related kinase (PIKK) and a novel polynucleotide encoding the same.

[0005]

[Means for Solving the Problems]

The present invention relates to (1) a polypeptide comprising an amino acid sequence consisting of 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, or (2) a polypeptide exhibiting an SMG-1 activity and comprising an amino acid sequence in which one or plural amino acids are deleted, substituted, and/or inserted at one or plural positions in an amino acid sequence consisting of 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2.

Further, the present invention relates to a polynucleotide encoding the polypeptide.

Further, the present invention relates to an expression vector comprising the polynucleotide.

Further, the present invention relates to a cell transfected with the expression vector.

Further, the present invention relates to an antibody which binds to the above polypeptide.

Further, the present invention relates to a method for screening a substance which inhibits an SMG-1 activity of the above polypeptide, comprising the steps of:
bringing into contact (1) the polypeptide, (2) Upf1/SMG-2, and (3) a substance to be tested; and
carrying out phosphorylation under the conditions that the polypeptide is brought into contact with Upf1/SMG-2 and the test substance, and analyzing whether or not Upf1/SMG-2 is phosphorylated.

[0006]

The term "SMG-1 activity" as used herein means an activity of phosphorylating Upf1/SMG-2 [Sun, X. et al., Proc. Natl. Acad. Sci. USA, 95, 10009-10014 (1998); and Bhattacharya, A. et al., RNA, 6, 1226-1235 (2000)].

[0007]

[Mode for Carrying out the Invention]

The present invention will be explained in detail hereinafter.

The present inventor found a novel PIKK consisting of 3657 amino acid residues, i.e., human SMG-1. The amino acid sequence thereof is the sequence consisting of the 1st to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2. Further, the present inventor found that a C-terminal fragment consisting of the 107th to 3657th amino acid residues in the novel protein and another C-terminal fragment consisting of the 129th to 3657th amino acid residues therein also exhibit an SMG-1 activity as well as the novel polypeptide. The present invention is based on these findings.

[0008]

The polypeptide of the present invention includes
(1) a polypeptide comprising the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2;
(2) a polypeptide exhibiting an SMG-1 activity and comprising an amino acid sequence in which one or plural amino acids are deleted, substituted, and/or inserted at one or plural positions in the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence

of SEQ ID NO: 2 (hereinafter referred to as a functionally equivalent mutant); and

(3) a polypeptide exhibiting an SMG-1 activity and comprising an amino acid sequence having a 90% or more homology, with the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, with the amino acid sequence consisting of the 1st to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, or with the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2 (hereinafter referred to as a homologous polypeptide).

[0009]

The "polypeptide comprising the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2 as the polypeptide of the present invention is not limited, so long as it is a polypeptide comprising the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity. It includes, for example,

(1a) a polypeptide having the base sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2;

(1b) a fusion polypeptide having an amino acid sequence in which an appropriate marker sequence or the like is added to the N-terminus and/or the C-terminus of the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity;

(1c) a polypeptide consisting of the amino acid sequence of SEQ ID NO: 2;

(1d) a fusion polypeptide having an amino acid sequence in which an appropriate marker sequence or the like is added to the N-terminus and/or the C-terminus of the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity;

(1e) a polypeptide having the base sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2; and

(1f) a fusion polypeptide having an amino acid sequence in which an appropriate marker sequence or the like is added to the N-terminus and/or the C-terminus of the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity.

[0010]

A method for confirming whether or not a polypeptide to be tested "exhibits an SMG-1 activity" as used herein is not particularly limited. It may be confirmed, for example, by carrying out phosphorylation under the conditions that the test polypeptide is brought into contact with Upf1/SMG-2 (for example, human Upf1/SMG-2), a fragment thereof capable of being phosphorylated, or a fusion polypeptide comprising Upf1/SMG-2 or the fragment thereof, and then analyzing whether or not Upf1/SMG-2, the fragment thereof, or the fusion polypeptide is phosphorylated, more particularly, for example, by the method described in Example 9(1).

[0011]

The above polypeptide (1a), i.e., "the polypeptide having the base sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2 is a novel protein consisting of 3551 amino acid residues and exhibiting an SMG-1 activity. The polypeptide (1a) corresponds to a partial polypeptide of the above polypeptide (1c), i.e., "the polypeptide consisting of the amino acid sequence of SEQ ID NO: 2 .

The polypeptide (1c) is a novel protein having a molecular weight of approximately 430 kDa, and referred to as "p430 in EXAMPLES.

The above polypeptide (1e), i.e., "the polypeptide having the base sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2 is a novel protein consisting of 3529 amino acid residues and exhibiting an SMG-1 activity. The polypeptide (1e) corresponds to a partial polypeptide of the polypeptide (1c), and is a novel protein having a molecular weight of approximately 400 kDa, and referred to as "p400 in EXAMPLES.

[0012]

As the marker sequence in the polypeptide of the present invention, for example, a sequence for easily carrying out confirmation of polypeptide expression, confirmation of intracellular localization thereof, purification thereof, or the like may be used. As the sequence, there may be mentioned, for example, the FLAG tag, the hexa-histidine tag, the hemagglutinin tag, the myc epitope, or the like.

[0013]

The functionally equivalent mutant of the present invention is not particularly limited, so long as it is a polypeptide comprising an amino acid sequence in which one or plural (preferably 1 to 10, more preferably 1 to 7, most preferably 1 to 5) amino acids, such as one to several amino acids, are deleted, substituted, and/or inserted at one or plural positions in the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity. Further, an origin of the functionally equivalent mutant is not limited to a human.

[0014]

The functionally equivalent mutant of the present invention includes, for example, human mutants of the polypeptide having the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and functionally equivalent mutants derived from organisms other than human (such as simian, mouse, rat, hamster, or dog). As the functionally equivalent mutants derived from organisms other than human, there may be mentioned, a simian native polypeptide having a molecular weight of 400 kDa or 430 kDa, a rat native polypeptide having a molecular weight of 400 kDa or 430 kDa, or a mouse native polypeptide having a molecular weight of 400 kDa or 430 kDa, as shown in Example 5.

Further, the functionally equivalent mutant of the present invention includes polypeptides prepared using polynucleotides obtained by artificially modifying polynucleotides encoding these native polypeptides (i.e., human mutants or functionally equivalent mutants derived from organisms other than human) or polynucleotides encoding

the polypeptide consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2 by genetic engineering techniques. The term "variation" as used herein means an individual difference between the same polypeptides in the same species or a difference between homologous polypeptides in several species.

[0015]

Human mutants of the polypeptide consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2 or functionally equivalent mutants derived from organisms other than a human may be obtained by those skilled in the art in accordance with the information of a base sequence (for example, the base sequence consisting of 712th to 11301st bases in the base sequence of SEQ ID NO: 1) of a polynucleotide encoding the polypeptide having the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2. In this connection, genetic engineering techniques may be generally performed in accordance with known methods (for example, Sambrook, J. et al., "Molecular Cloning-A Laboratory Manual", Cold Spring Harbor Laboratory, NY, 1989).

[0016]

For example, an appropriate probe or appropriate primers are designed in accordance with the information of a base sequence of a polynucleotide encoding the polypeptide having the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2. A polymerase chain reaction (PCR) method (Saiki, R. K. et al., Science, 239, 487-491, 1988) or a hybridization method is carried out using a sample (for example, total RNA or an mRNA fraction, a cDNA library, or a phage library) prepared from an organism (for example, a mammal such as human, simian, mouse, rat, hamster, or dog) of interest and the primers or the probe to obtain a polynucleotide encoding the polypeptide. A desired polypeptide may be obtained by expressing the resulting polynucleotide in an appropriate expression system and confirming that the expressed polypeptide exhibits an SMG-1 activity by, for example, the method described in Example 9(1).

[0017]

Further, the polypeptide artificially modified by genetic engineering techniques may be obtained by, for example, the following procedure. A gene encoding the polypeptide may be obtained by a conventional method, for example, site-directed mutagenesis (Mark, D. F. et al., Proc. Natl. Acad. Sci. USA, 81, 5662-5666, 1984). A desired polypeptide may be obtained by expressing the resulting polynucleotide in an appropriate expression system and confirming that the expressed polypeptide exhibits an SMG-1 activity by, for example, the method described in Example 9(1).

[0018]

The homologous polypeptide of the present invention is not particularly limited, so long as it is a polypeptide comprising an amino acid sequence having a 90% or more homology, with the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, with the amino acid sequence consisting of the 1st to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, or with the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity. The homologous polypeptide of the present invention may comprise an amino acid sequence having preferably a 95% or more homology, more preferably a 98% or more homology, most preferably a 99% or more homology, with respect to the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, the amino acid sequence consisting of the 1st to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, or the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2. As the homologous polypeptide of the present invention, a polypeptide having an amino acid sequence having a 90% or more homology (preferably a 95% or more homology, more preferably a 98% or more homology, most preferably a 99% or more homology), with the amino acid sequence consisting of the 129th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, with

the amino acid sequence consisting of the 1st to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, or with the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and exhibiting an SMG-1 activity is preferable.

The term "homology" as used herein means a value obtained by BLAST [Basic local alignment search tool; Altschul, S. F. et al., J. Mol. Biol., 215, 403-410, (1990)].

[0019]

Further, the polypeptide of the present invention includes a polypeptide obtained by bringing mammalian cells or disrupted cells (such as cell lysate) into contact with an antibody specific for SMG-1 to form an immunocomplex (such as immunoprecipitate) and then removing the antibody from the immunocomplex. As the polypeptide, there may be mentioned, for example, a human, simian, rat, or mouse native polypeptide having a molecular weight of 400 kDa or 430 kDa.

[0020]

The polynucleotide of the present invention is not particularly limited, so long as it encodes the polypeptide of the present invention. As the polynucleotide of the present invention, there may be mentioned, for example, a polynucleotide comprising the base sequence consisting of the 712th to 11301st bases in the base sequence of SEQ ID NO: 1, and

(i) the polynucleotide having the base sequence consisting of the 646th to 11301st bases in the base sequence of SEQ ID NO: 1 [and encoding the above polypeptide (1a) of the present invention];

(ii) the polynucleotide having the base sequence consisting of the 328th to 11301st bases in the base sequence of SEQ ID NO: 1 [and encoding the above polypeptide (1c) of the present invention]; or

(iii) the polynucleotide having the base sequence consisting of the 712th to 11301st bases in the base sequence of SEQ ID NO: 1 [and encoding the above polypeptide (1e) of the present invention]

is preferable. In this connection, the term "polynucleotide" as used herein includes both DNA and RNA.

[0021]

A method for producing the polynucleotide of the present invention is not particularly limited, but there may be mentioned, for example, (1) a method using PCR, (2) a method using conventional genetic engineering techniques (i.e., a method for selecting a transformant comprising a desired cDNA from strains transformed with a cDNA library), or (3) a chemical synthesis method. These methods will be explained in this order hereinafter.

[0022]

In the method using PCR of the item (1), the polynucleotide of the present invention may be produced, for example, by the following procedure.

mRNA is extracted from human cells or tissue capable of producing the polypeptide of the present invention. A pair of primers, between which full-length mRNA corresponding to the polypeptide of the present invention or a partial region of the mRNA is located, is synthesized on the basis of the base sequence of a polynucleotide encoding the polynucleotide of the present invention. Full-length cDNA encoding the polypeptide of the present invention or a part of the cDNA may be obtained by performing a reverse transcriptase-polymerase chain reaction (RT-PCR) using the extracted mRNA as a template.

[0023]

More particularly, total RNA containing mRNA encoding the polypeptide of the present invention is extracted by a known method from cells or tissue capable of producing the polypeptide of the present invention. As an extraction method, there may be mentioned, for example, a guanidine thiocyanate-hot phenol method, a guanidine thiocyanate-guanidine hydrochloride method, or a guanidine thiocyanate-cesium chloride method. The guanidine thiocyanate-cesium chloride method is preferably used. The cells or tissue capable of producing the polypeptide of the present invention may be identified, for example, by a northern blotting method using a polynucleotide or a part thereof

encoding the polypeptide of the present invention or a western blotting method using an antibody specific for the polypeptide of the present invention.

[0024]

Next, the extracted mRNA is purified. Purification of the mRNA may be made in accordance with a conventional method. For example, the mRNA may be purified by adsorption and elution using an oligo(dT)-cellulose column. The mRNA may be further fractionated by, for example, a sucrose density gradient centrifugation, if necessary. Alternatively, commercially available extracted and purified mRNA may be used without carrying out the extraction of the mRNA.

Next, the first-strand cDNA is synthesized by carrying out a reverse transcriptase reaction of the purified mRNA in the presence of a random primer, an oligo dT primer, and/or a custom primer. This synthesis may be carried out in accordance with a conventional method. The resulting first-strand cDNA is subjected to PCR using two primers between which a full-length or a partial region of the polynucleotide of interest is located, thereby amplifying the cDNA of interest. The resulting DNA is fractionated by, for example, an agarose gel electrophoresis. The DNA fragment of interest may be obtained by carrying out a digestion of the DNA with restriction enzymes and subsequent ligation, if necessary.

[0025]

In the method using conventional genetic engineering techniques of the item (2), the polynucleotide of the present invention may be produced, for example, by the following procedure.

First, single-stranded cDNA is synthesized by using reverse transcriptase from mRNA prepared by the above-mentioned PCR method as a template, and then double-stranded cDNA is synthesized from the single-stranded cDNA. As this method, there may be mentioned, for example, an S1 nuclease method (Efstratiadis, A. et al., Cell, 7, 279-288, 1976), a Land method (Land, H. et al., Nucleic Acids Res., 9, 2251-2266, 1981), an O. Joon Yoo method (Yoo, O. J. et al., Proc.

Natl. Acad. Sci. USA, 79, 1049-1053, 1983), and an Okayama-Berg method (Okayama, H. and Berg, P., Mol. Cell. Biol., 2, 161-170, 1982).

[0026]

Next, a recombinant plasmid comprising the double-stranded cDNA is prepared and introduced into an Escherichia coli strain, such as DH 5α, HB101, or JM109, thereby transforming the strain. A transformant is selected using a drug resistance against, for example, tetracycline, ampicillin, or kanamycin as a marker. When the host cell is E. coli, transformation of the host cell may be carried out, for example, by the method of Hanahan (Hanahan, D. J., Mol. Biol., 166, 557-580, 1983); namely, a method in which the recombinant DNA is added to competent cells prepared in the presence of $CaCl_2$, $MgCl_2$, or RbCl. Further, as a vector other than a plasmid, a phage vector such as a lambda system may be used.

[0027]

As a method for selecting a transformant containing the cDNA of interest from the resulting transformants, various methods such as (i) a method for screening a transformant using a synthetic oligonucleotide probe, (ii) a method for screening a transformant using a probe produced by PCR, (iii) a method for screening a transformant using an antibody against the polypeptide of the present invention, or (iv) a method for screening a transformant using a selective hybridization translation system, may be used.

[0028]

In the method of the item (i) for screening a transformant using a synthetic oligonucleotide probe, the transformant containing the cDNA of interest may be selected, for example, by the following procedure.

An oligonucleotide which corresponds to the whole or a part of the polypeptide of the present invention is synthesized (in this case, it may be either a nucleotide sequence taking the codon usage into consideration or a plurality of nucleotide sequences as a combination of possible nucleotide sequences, and in the latter case, their numbers can be reduced by including inosine) and, using this

oligonucleotide as a probe (labeled with ^{32}P or ^{33}P), hybridized with a nitrocellulose filter or a polyamide filter on which DNAs of the transformants are denatured and fixed, to screen and select resulting positive strains.

[0029]

In the method of the item (ii) for screening a transformant using a probe produced by PCR, the transformant containing the cDNA of interest may be selected, for example, by the following procedure.

Oligonucleotides of a sense primer and an antisense primer corresponding to a part of the polypeptide of the present invention are synthesized, and a DNA fragment encoding the whole or a part of the polypeptide of interest is amplified by carrying out PCR using these primers in combination. As a template DNA used in this method, cDNA synthesized by a reverse transcription reaction from mRNA of cells capable of producing the polypeptide of the present invention, or genomic DNA, may be used. The resulting DNA fragment is labeled with ^{32}P or ^{33}P, and a transformant containing the cDNA of interest is selected by carrying out a colony hybridization or a plaque hybridization using this fragment as a probe.

[0030]

In the method of the item (iii) for screening a transformant using an antibody against the polypeptide of the present invention, the transformant containing the cDNA of interest may be selected, for example, by the following procedure.

Polypeptides are produced into a culture supernatant, inside the cells, or on the cell surface of transformants. A transformant containing the cDNA of interest is selected by detecting a strain producing the desired polypeptide using an antibody against the polypeptide of the present invention and a second antibody against the first antibody.

[0031]

In the method of the item (iv) for screening a transformant using a selective hybridization translation system, the transformant containing the cDNA of interest may be selected, for example, by the following procedure.

First, cDNA obtained from each transformant is blotted on, for example, a nitrocellulose filter and hybridized with mRNA prepared from cells capable of producing the polypeptide of the present invention, and then the mRNA bound to the cDNA is dissociated and recovered. The recovered mRNA is translated into a polypeptide in an appropriate polypeptide translation system, for example, injection into Xenopus oocytes or a cell-free system such as a rabbit reticulocyte lysate or a wheat germ. A transformant containing the cDNA of interest is selected by detecting it with the use of an antibody against the polypeptide of the present invention.

[0032]

A method for collecting the polynucleotide of the present invention from the resulting transformant of interest can be carried out in accordance with a known method (for example, Sambrook, J. et al., "Molecular Cloning-A Laboratory Manual", Cold Spring Harbor Laboratory, NY, 1989). For example, it may be carried out by separating a fraction corresponding to the plasmid DNA from cells and cutting out the cDNA region from the plasmid DNA.

[0033]

In the chemical synthesis method of the item (3), the polynucleotide of the present invention may be produced, for example, by binding DNA fragments produced by a chemical synthesis method. Each DNA can be synthesized using a DNA synthesizer [for example, Oligo 1000M DNA Synthesizer (Beckman) or 394 DNA/RNA Synthesizer (Applied Biosystems)].

Further, the polynucleotide of the present invention may be produced by nucleic acid chemical synthesis in accordance with a conventional method such as a phosphite triester method (Hunkapiller, M. et al., Nature, 10, 105-111, 1984), based on the information on the polypeptide of the present invention. In this connection, codons for each amino acid are known and can be optionally selected and determined by the conventional method, for example, by taking a codon usage of each host to be used into consideration (Crantham, R. et al., Nucleic Acids Res., 9, r43-r74, 1981). Further, a partial modification of codons of these base sequences can

be carried out in accordance with a conventional method, such as site directed mutagenesis which uses a primer comprised of a synthetic oligonucleotide coding for a desired modification (Mark, D. F. et al., Proc. Natl. Acad. Sci. USA, 81, 5662-5666, 1984).

[0034]

Determination of the DNA sequences obtained by the above-mentioned methods can be carried out by, for example, a Maxam-Gilbert chemical modification method (Maxam, A. M. and Gilbert, W., "Methods in Enzymology", 65, 499-559, 1980) or a dideoxynucleotide chain termination method (Messing, J. and Vieira, J., Gene, 19, 269-276, 1982).

[0035]

An isolated polynucleotide of the present invention is re-integrated into an appropriate vector DNA and a eucaryotic or procaryotic host cell may be transfected by the resulting expression vector. Further, it is possible to express the polynucleotide in a desired host cell, by introducing an appropriate promoter and a sequence related to the gene expression into the vector.

[0036]

The expression vector of the present invention is not particularly limited, so long as it comprises the polynucleotide of the present invention. As the expression vector, there may be mentioned, for example, an expression vector obtained by introducing the polynucleotide of the present invention into a known expression vector appropriately selected in accordance with a host cell to be used or a cell to be introduced.

[0037]

The cell of the present invention is not particularly limited, so long as it is transfected with the expression vector of the present invention and comprises the polynucleotide of the present invention. The cell of the present invention may be, for example, a cell in which the polynucleotide is integrated into a chromosome of a host cell, or a cell containing the polynucleotide as an expression vector comprising polynucleotide. Further, the cell of the present invention may be a cell expressing the

polypeptide of the present invention, or a cell not expressing the polypeptide of the present invention. The cell of the present invention may be obtained by, for example, transfecting a desired host cell with the expression vector of the present invention.

[0038]

In the eucaryotic host cells, for example, cells of vertebrates, insects, and yeast are included. As the vertebral cell, there may be mentioned, for example, a simian COS cell (Gluzman, Y., Cell, 23, 175-182, 1981), a dihydrofolate reductase defective strain of a Chinese hamster ovary cell (CHO) (Urlaub, G. and Chasin, L. A., Proc. Natl. Acad. Sci. USA, 77, 4216-4220, 1980), a human fetal kidney derived HEK293 cell, a 293-EBNA cell (Invitrogen) obtained by introducing an EBNA-1 gene of Epstein Barr Virus into HEK293 cell, or a human 293T cell (DuBridge, R. B. et al., Mol. Cell. Biol., 7, 379-387, 1987).

[0039]

As an expression vector for a vertebral cell, a vector containing a promoter positioned upstream of the gene to be expressed, an RNA splicing site, a polyadenylation site, a transcription termination sequence, and the like may be generally used. The vector may further contain a replication origin, if necessary. As the expression vector, there may be mentioned, for example, pSV2dhfr containing an SV40 early promoter (Subramani, S. et al., Mol. Cell. Biol., 1, 854-864, 1981), pEF-BOS containing a human elongation factor promoter (Mizushima, S. and Nagata, S., Nucleic Acids Res., 18,5322, 1990), or pCEP4 containing a cytomegalovirus promoter (Invitrogen).

[0040]

When the COS cell is used as the host cell, a vector which has an SV40 replication origin, can perform an autonomous replication in the COS cell, and has a transcription promoter, a transcription termination signal, and an RNA splicing site, may be used as the expression vector. As the vector, there may be mentioned, for example, pME18S (Maruyama, K. and Takebe, Y., Med. Immunol., 20, 27-

32, 1990), pEF-BOS (Mizushima, S. and Nagata, S., Nucleic Acids Res., 18, 5322, 1990), or pCDM8 (Seed, B., Nature, 329, 840-842, 1987).

[0041]

The expression vector may be incorporated into COS cells by, for example, a DEAE-dextran method (Luthman, H. and Magnusson, G., Nucleic Acids Res., 11, 1295-1308, 1983), a calcium phosphate-DNA co-precipitation method (Graham, F. L. and van der Ed, A. J., Virology, 52, 456-457, 1973), a method using a commercially available transfection reagent (for example, FuGENE™6 Transfection Reagent; Boeringer Mannheim), or an electroporation method (Neumann, E. et al., EMBO J., 1, 841-845, 1982).

[0042]

When the CHO cell is used as the host cell, a transfected cell capable of stably producing the polypeptide of the present invention can be obtained by carrying out co-transfection of an expression vector comprising the polynucleotide encoding the polypeptide of the present invention, together with a vector capable of expressing a neo gene which functions as a G418 resistance marker, such as pRSVneo (Sambrook, J. et al., "Molecular Cloning-A Laboratory Manual", Cold Spring Harbor Laboratory, NY, 1989) or pSV2-neo (Southern, P. J. and Berg, P., J. Mol. Appl. Genet., 1, 327-341,1982), and selecting a G418 resistant colony.

[0043]

The cell of the present invention may be cultured in accordance with the conventional method, and the polypeptide of the present invention is produced inside the cells. As a medium to be used in the culturing, a medium commonly used in a desired host cell may be appropriately selected. In the case of the COS cell, for example, a medium such as an RPMI-1640 medium or a Dulbecco's modified Eagle's minimum essential medium (DMEM) may be used, by supplementing it with a serum component such as fetal bovine serum (FBS) if necessary. In the case of the 293-EBNA cell, a medium such as a Dulbecco's modified Eagle's minimum essential medium (DMEM) with a serum component such as fetal bovine serum

(FBS) and G418 may be used.

[0044]

The polypeptide of the present invention produced inside the cell of the present invention by culturing the cells may be separated and purified therefrom by various known separation techniques making use of the physical properties, chemical properties and the like of the polypeptide. More particularly, the polypeptide of the present invention may be purified by treating a cell extract containing the polypeptide of the present invention with a commonly used treatment, for example, a treatment with a protein precipitant, ultrafiltration, various liquid chromatography techniques such as molecular sieve chromatography (gel filtration), adsorption chromatography, ion exchange chromatography, affinity chromatography, or high performance liquid chromatography (HPLC), or dialysis, or a combination thereof.

[0045]

When the polypeptide of the present invention is expressed as a fusion protein with a marker sequence in frame, identification of the expression of the polypeptide of the present invention, purification thereof, or the like may be easily carried out. As the marker sequence, there may be mentioned, for example, a FLAG tag, a hexa-histidine tag, a hemagglutinin tag, or a myc epitope. Further, by inserting a specific amino acid sequence recognized by a protease such as enterokinase, factor Xa, or thrombin between the marker sequence and the polypeptide of the present invention, the marker sequence may be removed by the protease.

[0046]

It is possible to screen a substance which modifies (for example, inhibits or promotes) an SMG-1 activity of the polypeptide according to the present invention, using the polypeptide of the present invention.

A substance inhibiting the SMG-1 activity of the polypeptide of the present invention (for example, an inhibitor of phosphatidyl inositol kinase related kinase, more particularly, for example, wortmannin or caffeine) can

suppress NMD, and thus is useful as a candidate of an agent for treating and/or preventing a disease caused by at least a premature translation termination codon (PTC) generated by a nonsense mutation. The polypeptide of the present invention per se may be used as a screening tool for screening a substance inhibiting the SMG-1 activity of the polypeptide of the present invention, or for screening an agent for treating and/or preventing a disease caused by a nonsense mutation of a specific gene. The disease caused by one or more PTCs generated by a nonsense mutation is not particularly limited, but there may be mentioned, for example, a genetic disease (for example, Duchenne type muscular dystrophy), cancer due to a somatic mutation, or the like. The important point is that, among all diseases due to genome mutation, almost all diseases "due to one or more PTCs by a nonsense mutation" are included in such diseases.

[0047]

One-quarter of the diseases due to genome mutations have the termination codon in the middle of a specific gene. The reasons for these diseases are that the protein consisting of the full-length polypeptide inherently encoded by the gene is not expressed, and that, due to the presence of the NMD mechanism, almost no protein fragments consisting of the N terminal side partial fragments of the full length polypeptide inherently encoded by the gene are expressed. However, even if there is a termination codon in the middle of the gene, and even if in the state of a protein fragment, there are not a few cases of activity of the same extent as that of full length polypeptide or the minimum necessary level, depending on the type of the gene or the position of the termination codon. In this case, if it were possible to inhibit the NMD mechanism, it would become possible to express a protein fragment having an effective activity, and thus it is theoretically predicted that at least part of a disease due to the presence of a termination codon in the middle of a specific gene, that is, a disease due to nonsense mutation of a specific gene can be alleviated. However, no technique for a specific suppression of NMD has

been known at all in the past.

Among the substances selected by the screening method of the present invention, a substance inhibiting the SMG-1 activity of the polypeptide of the present invention can specifically suppress NMD through inhibition of the SMG-1 activity of the polypeptide of the present invention, and thus is useful as an active ingredient of a new type of agent for treatment and/or prevention which can alleviate gene mutations for at least part of all sorts of diseases due to the nonsense mutation of specific genes.

[0048]

The screening method of the present invention comprises the steps of:

bringing into contact (1) the polypeptide of the present invention, (2) Upf1/SMG-2 (for example, human Upf1/SMG-2), and (3) a substance to be tested; and

carrying out phosphorylation under the conditions that the polypeptide is brought into contact with Upf1/SMG-2 and the test substance, and analyzing whether or not Upf1/SMG-2 is phosphorylated.

[0049]

Substances to be tested which may be applied to the detection method or screening method of the present invention are not particularly limited, but there may be mentioned, for example, various known compounds (including peptides) registered in chemical files, compounds obtained by combinatorial chemistry techniques (Terrett, N. K. et al., Tetrahedron, 51, 8135-8137, 1995) or conventional synthesis techniques, or random peptides prepared by employing a phage display method (Felici, F. et al., J. Mol. Biol., 222, 301-310, 1991) or the like. In addition, culture supernatants of microorganisms, natural components derived from plants or marine organisms, or animal tissue extracts may be used as the test Substances for screening. Further, compounds (including peptides) obtained by chemically or biologically modifying compounds (including peptides) selected by the screening method of the present invention may be used.

[0050]

The screening method of the present invention can be performed in the same way as the above-mentioned method of judgment of the SMG-1 activity, except that, instead of bringing the test polypeptide into contact with Upf1/SMG-2, the polypeptide of the present invention, Upf1/SMG-2, and the test substance are brought into contact. That is, it is possible to judge whether or not the test substance inhibits the SMG-1 activity of the polypeptide of the present invention, by bringing into contact the polypeptide of the present invention, Upf1/SMG-2, and the test substance, carrying out phosphorylation in the presence of the test substance, and then analyzing whether or not Upf1/SMG-2 is phosphorylated. When the Upf1/SMG-2 is not phosphorylated or the degree of the phosphorylation thereof decreases in the presence of the test substance, it is possible to judge that the test substance is a substance inhibiting the SMG-1 activity of the polypeptide of the present invention.

[0051]

An antibody, such as a polyclonal antibody or a monoclonal antibody, which reacts with the polypeptide of the present invention may be obtained by directly administering the polypeptide of the present invention or a fragment thereof to various animals. Alternatively, it may be obtained by a DNA vaccine method (Raz, E. et al., Proc. Natl. Acad. Sci. USA, 91, 9519-9523, 1994; or Donnelly, J. J. et al., J. Infect. Dis., 173, 314-320, 1996), using a plasmid into which a polynucleotide encoding the polypeptide of the present invention is inserted.

[0052]

The polyclonal antibody may be produced from a serum or eggs of an animal such as a rabbit, a rat, a goat, or a chicken, in which the animal is immunized and sensitized by the polypeptide of the present invention or a fragment thereof emulsified in an appropriate adjuvant (for example, Freund's complete adjuvant) by intraperitoneal, subcutaneous, or intravenous administration. The polyclonal antibody may be separated and purified from the resulting serum or eggs in accordance with conventional methods for polypeptide isolation and purification. Examples of the

separation and purification methods include, for example, centrifugal separation, dialysis, salting-out with ammonium sulfate, or a chromatographic technique using such as DEAE-cellulose, hydroxyapatite, protein A agarose, and the like.

[0053]

The monoclonal antibody may be easily produced by those skilled in the art, according to, for example, a cell fusion method of Kohler and Milstein (Kohler, G. and Milstein, C., Nature, 256, 495-497, 1975).

A mouse is immunized intraperitoneally, subcutaneously, or intravenously several times at an interval of a few weeks by a repeated inoculation of emulsions in which the polypeptide of the present invention or a fragment thereof is emulsified into a suitable adjuvant such as Freund's complete adjuvant. Spleen cells are removed after the final immunization, and then fused with myeloma cells to prepare hybridomas.

[0054]

As a myeloma cell for obtaining a hybridoma, a myeloma cell having a marker such as a deficiency in hypoxanthine-guanine phosphoribosyltransferase or thymidine kinase (for example, mouse myeloma cell line P3X63Ag8.U1) may be used. As a fusing agent, polyethylene glycol may be used. As a medium for preparation of hybridomas, for example, a commonly used medium such as an Eagle's minimum essential medium, a Dulbecco's modified minimum essential medium, or an RPMI-1640 medium may be used by adding properly 10 to 30% of a fetal bovine serum. The fused strains may be selected by a HAT selection method. A culture supernatant of the hybridomas is screened by a well-known method such as an ELISA method or an immunohistological method, to select hybridoma clones secreting the antibody of interest. The monoclonality of the selected hybridoma is guaranteed by repeating subcloning by a limiting dilution method. Antibodies in an amount which may be purified are produced by culturing the resulting hybridomas in a medium for 2 to 4 days, or in the peritoneal cavity of a pristane-pretreated BALB/c strain mouse for 10 to 20 days.

[0055]

The resulting monoclonal antibodies in the culture supernatant or the ascites may be separated and purified by conventional polypeptide isolation and purification methods. Examples of the separation and purification methods include, for example, centrifugal separation, dialysis, salting-out with ammonium sulfate, or chromatographic technique using such as DEAE-cellulose, hydroxyapatite, protein A agarose, and the like.

Further, the monoclonal antibodies or the antibody fragments containing a part thereof may be produced by inserting the whole or a part of a gene encoding the monoclonal antibody into an expression vector and introducing the resulting expression vector into appropriate host cells (such as E. coli, yeast, or animal cells).

[0056]

Antibody fragments comprising an active part of the antibody such as $F(ab')_2$, Fab, Fab', or Fv may be obtained by a conventional method, for example, by digesting the separated and purified antibodies (including polyclonal antibodies and monoclonal antibodies) with a protease such as pepsin or papain, and separating and purifying the resulting fragments by standard polypeptide isolation and purification methods.

[0057]

Further, an antibody which reacts to the polypeptide of the present invention may be obtained in a form of single chain Fv or Fab in accordance with a method of Clackson et al. or a method of Zebedee et al. (Clackson, T. et al., Nature, 352, 624-628, 1991; or Zebedee, S. et al., Proc. Natl. Acad. Sci. USA, 89, 3175-3179, 1992). Furthermore, a humanized antibody may be obtained by immunizing a transgenic mouse in which mouse antibody genes are substituted with human antibody genes (Lonberg, N. et al., Nature, 368, 856-859, 1994).

[0058]

[EXAMPLES]

The present invention now will be further illustrated by, but is by no means limited to, the following Examples.

Example 1: Cloning of Human SMG-1 (hSMG-1) cDNA

The present inventor discovered that the N-terminus of the amino acid sequence encoded by the human cDNA clone KIAA0421 [Ishikawa, K. et al., DNA Res., 4, 307 (1997); GenBank access no. AB007881] has homology with the amino acid sequence characteristic of the kinase domain conserved in the PIKK family, and that the C-terminus has homology with the amino acid sequence characteristic of the FAT domain conserved in the PIKK family [Bosotti et al., Trends Biochem. Sci., 25, 225 (2000)]. Therefore, the human cDNA clone KIAA0421 was considered to be a novel cDNA of the PIKK family, but while this base sequence includes a termination codon and 3 nontranslation region, there is no sequence capable of being specified as the start codon, and thus it was considered that the cDNA was of incomplete length. Therefore, to clarify the base sequence of the full-length cDNA, it was attempted to obtain the further 5 side cDNA clone from the clone KIAA0421.

[0059]

Using a cDNA fragment of the human cDNA clone KIAA0421 as a probe, a clone C was isolated from a cDNA library of the human cell line HeLa (Clonetech). Similarly, a clone yama9 (Y9) was isolated from a HeLa cDNA library [Chambon et al., Proc. Natl. Acad. Sci. USA, 86 (14), 5310-5314], a clone liver33 (Liv33) was isolated from a human liver library (Clonetech), and a clone muscle29 (mus29) was isolated from a human muscle library (Clonetech). Further, other various clones were isolated. The base sequences thereof were determined.

[0060]

Next, a combination of a forward primer consisting of the base sequence of SEQ ID NO: 3 and a reverse primer consisting of the base sequence of SEQ ID NO: 4 was used to obtain a clone gap1 by a reverse transcription polymerase chain reaction (RT-PCR) method using the Total RNA of the human cell line HeLa. The RT-PCR was performed by using a commercially available kit (Ready-To-Go RT-PCR beads; Pharmacia), and performing an RT reaction at 42°C for 30 minutes, then performing heat denaturation at 95°C (3 minutes), repeating a cycle of 95°C (1 minute), 54°C (1

minute), and 72°C (1 minute) 32 times, and finally performing an elongation reaction at 72°C (7 minutes).

Further, a combination of a forward primer consisting of the base sequence of SEQ ID NO: 5 and a reverse primer consisting of the base sequence of SEQ ID NO: 6 was used to obtain a clone gap2 by the RT-PCR method using the Total RNA of the human cell line HeLa. The RT-PCR was performed under the same conditions as the RT-PCR when obtaining the clone gap1.

It was attempted to connect the base sequences of these clones, but there was no sequence capable of being specified as the start codon, and only a base sequence of cDNA of an incomplete length could be obtained.

[0061]

Therefore, a search for an EST having a sequence matching with the obtained base sequence was made in the base sequence database (GenBank), whereupon the human EST clone AI005513 (Research Genetics) was found. The base sequence of this EST has a start codon ATG in its frame, so the EST of the region including the start coden of the full-length cDNA consisting of the human cDNA clone KIAA0421 and its upstream region was estimated.

By determining the base sequence of the human EST clone AI005513, the base sequence of the cDNA consisting of the human cDNA clone KIAA0421 and its upstream region was clarified. The base sequence was that of SEQ ID NO: 1. When the base sequence database (GenBank) was searched, it was found that this base sequence was novel.

[0062]

The relationship between the obtained cDNA clones and the novel base sequences and open reading frame (ORF) obtained therefrom is shown in Fig. 1. The length of the cDNA consisting of KIAA0421 and its upstream region, obtained from each cDNA clone, was approximately 13 kb. There was an approximately 11 kb open reading frame (ORF) encoding a protein consisting of 3657 amino acids. The estimated molecular weight of the protein encoded by the ORF was approximately 430 kDa, which matched the roughly calculated molecular weight of the endogenous molecule

(p430) detected in Example 5(1).

[0063]

A search of homology was conducted for the amino acid sequence (amino acid sequence of SEQ ID NO: 2) encoded by the ORF, whereupon it was found that there was a homology with the PIKK family FRAP (FKBP12-rapamycin associated protein)/mTOR (mammalian target of rapamycin)/RAFT1 (rapamycin and FKBP-target 1), ATM (ataxia telangiectasia mutated), ATR (ATM- and Rad3-related)/FRAP1, DNA-PKcs (DNA-PK catalytic subunit) and the like. The results of a comparison of human SMG-1 and known proteins are shown in Fig. 2.

[0064]

In Fig. 2, the deduced PIKK related domain is shown by the black square. The FKBP12/rapamycin binding region (FRB) and its homologous region (FRBH) is shown by the dark gray, and the RAD3 homologous region is shown by the light gray. CR1 to CR6 mean regions with a high homology with C. elegans SMG1 (CeSMG1), and "1000 a.a." shows the length of 1000 amino acid residues. Further, the numerical values of the homology are from GeneWorks ver 2.5.1 (IntelliGenetics). GenBank access number of FRAP is L34075, that of ATM is U33841, that of ATR is U76308, and that of DNA-PKcs is U34994.

[0065]

In human SMG-1, the CR1 is the region consisting of the 557th to 727th amino acids. Similarly, the CR2 is the region consisting of the 911st to 1051st amino acids, the CR3 is the region consisting of the 1560th to 1756th amino acids, the CR4 is the region consisting of the 1785th to 2107th amino acids, the CR5 is the region consisting of the 2141st to 2422nd amino acids, and the CR6 is the region consisting of the 3602nd to 3657th amino acids.

Further, the region consisting of the 2130th to 2136th amino acids in the human SMG-1 is an amino acid sequence capable of functioning as an NLS (nuclear localization signal).

[0066]

Further, a molecular phylogenetic tree for the obtained

novel sequence and the PIKK family molecules was prepared on the basis of the amino acid sequences, whereupon the cDNA consisting of the human cDNA clone KIAA0421 and its upstream region is closest to fruit-fly SMG-1 and C. elegans SMG-1, which are genes involved in the degradation of abnormal RNA, and thus was considered to encode human SMG-1. In this connection, human SMG-1 includes a sequence FRBH (FKBP12/rapamycin binding homology) having homology with the FKBP12/rapamycin binding site of FRAP/mTOR/RAFT1. Further, unlike other PIKK families, a long sequence of an approximately 1200 amino acids was inserted between the kinase domain and the FAT domain.

[0067]

Example 2: Detection of mRNA of Human SMG-1 in Various Human Cell Lines by Northern Blotting

A total RNA was prepared from human cell lines HPB-ALL [Morikawa, S. et al., Int. J. Cancer, 21, 166 (1978)], HL-60 (CCL-240), U937 [Sundstrom, C. et al., Int. J. Cancer, 17, 565 (1976)], HepG2 (HB-8065), HeLa (CCL-2), PC3, A498, and 5873T using an RNA extraction kit (Quick Prep Total RNA extraction kit; Amersham Pharmacia Biotech) in accordance with the manual attached to the kit. The following blotting and hybrizing were performed in accordance with the document [Sugiyama, JBC, 275, 1095-1104, (2000)]. More particularly, the RNAs were electrophoresed, and then transferred to a polyamide membrane (Hybond; Amersham Pharmacia Biotech). The 5 -side fragment (corresponding to the base sequence consisting of the 6255th to 7048th bases in the base sequence of SEQ ID NO: 1) of the cDNA clone KIAA0421 of human SMG-1 was labeled using a Multiprime DNA Labelling System (Amersham Pharmacia Biotech) in accordance with the manual attached to the kit and using [α-^{32}P]dCTP (220 TBq/mmol; Amersham Pharmacia Biotech). The polyamide membrane to which the RNA has been transferred was hybridized with the labeled cDNA fragment as a probe, and was washed with 0.1×SSC [1.67 mmol/L sodium chloride and 1.67 mmol/L sodium citrate (pH7.0)]-0.1% sodium dodecyl sulfate (SDS) at 60°C (30 minutes) three times, and then the signal was detected by autoradiography.

[0068]

The results of autoradiography for HPB-ALL, U937, HepG2, HeLa, and PC3 are shown in Fig. 3. In Fig. 3, "28S" and "18S" show the electrophoresis positions of the 28S libosome RNA and 18S libosome RNA, respectively. As shown in Fig. 3, the two bands of mRNA of human SMG-1 shown by the arrows were detected. Further, in all remaining human cell lines (A549 and 293T), two bands were similarly detected (data not shown). Therefore, it was considered that two types of lengths of mRNAs were transcribed from the human SMG-1 gene.

[0069]

Example 3: Mapping of Human Chromosome by Fluorescent In Situ Hybridization (FISH) Method

FISH mapping was performed in accordance with the document [Izumi et al., JCB, 143, 95-106 (1998)]. More particularly, lymphocytes isolated from human blood were cultured, using a medium MEM (Minimal Essential Medium) to which 10% fetal bovine serum and phytohemagglutinin were added, at 37°C for 68 to 72 hours. To the lymphocytes cultured while synchronizing the cell cycle, 0.18 mg/mL bromodeoxyuridine (BrdU; Sigma Aldrich) was added to be incorporated into the cells. The cells were washed three times with a serum-free medium, and then were recultured using an MEM containing 2.5 mg/mL thymidine (Sigma Aldrich) at 37°C for 6 hours. The cells were collected and a slide was prepared by the standard method of a hyposmotic treatment, fixation, and air drying.

[0070]

As the FISH probe, the cDNA clone KIAA0421 of human SMG-1 (full-length) was biotinylated using biotinylated dATP and a BioNick Labelling Kit (Life Technologies) at 15°C for 1 hour [Heng HH et al., Proc. Natl. Acad. Sci. USA, 89, 9509-9513 (1992)]. In situ hybridization and its detection were performed in accordance with the method of the documents [Heng HH et al., Proc. Natl. Acad. Sci. USA, 89, 9509 (1992); Heng HH and Tsui LC, Chromosoma, 102, 325 (1993)]. Simply explained, the slide was heated at 55°C for 1 hour (i.e., a ribonuclease treatment), then the slide was treated at 70°C for 2 minutes using 2xSSC [33.3 mmol/L sodium

chloride and 33.3 mmol/L sodium citrate (pH7.0)] containing 70% formaldehyde to denature the chromosomes, and dehydrated by ethanol. The probe was placed on the slide of the denatured chromosomes to perform hybridization overnight, and then the slide was washed and applied to the detection system. A signal appeared on the 16th chromosome, whereby it was found that the human SMG-1 gene is located on the 16th chromosome (16p12).

[0071]

Example 4: Preparation of Antibody for Human SMG-1

Anti-human SMG-1 antiserum P1, antiserum C3, antiserum L1, antiserum L2, antiserum N1, and antiserum N2 were prepared by immunizing rabbits (New Zealand White) using the following immunogen together with adjuvants. As the adjuvants, Titer Max Gold (CytRx) was used for antiserum LT and antiserum NT, and Freund's adjuvant (Wako Pure Chemicals) was used for antisera other than antiserum LT and antiserum NT.

[0072]

As the immunogen for antiserum P1, a peptide consisting of 15 amino acids corresponding to the C-terminus of human SMG-1 and bonded with keyhole limpet hemocyanin (KLH) was used. The peptide has an amino acid sequence wherein the cysteine residue was added to the N-terminus of the amino acid sequence of SEQ ID NO: 7 (CDNLAQLYEGWTAWV; i.e., the sequence consisting of the 3644th to 3657th amino acid residues in the amino acid sequence of SEQ ID NO: 2).

To prepare antiserum C3, a 1.4kb MscI-MscI fragment (corresponding to the base sequence consisting of the 7641st to 9186th bases in the base sequence of SEQ ID NO: 1, and covering a half of the kinase insertion region at the C-terminal side) of the human SMG-1 cDNA of clone KIAA0421 was inserted into the SmaI site of the vector pGEX6P-3 (Amersham Pharmacia Biotech) for expressing a fusion protein with glutathione S-transferase (GST). E. coli BL21 was transformed with the plasmid to express the C-terminal fragment [corresponding to the amino acid sequence consisting of the 3076th to 3542nd amino acid residues in the human SMG-1 amino acid sequence (amino acid sequence of

SEQ ID NO: 2)] of human SMG-1, as a fusion protein (molecular weight = approximately 70 kDa) with GST. The fusion protein produced in E. coli formed insoluble inclusion bodies. The purified inclusion bodies were dissolved in 1×SDS sample buffer [100 mmol/L TrisHCl (pH6.8), 2% SDS, 6% β-mercaptoethanol (β-ME), 10% glycerol, and 0.01% Bromophenol Blue]. SDS polyacryl amide gel electrophoresis (SDS-PAGE) was performed, and then the 70 kDa protein band was cut from the gel, finely pulverized, and used as the immunogen.

[0073]

To prepare antiserum L1 and antiserum L2, similarly as the case of antiserum C3, an approximately 600bp of cDNA fragment (corresponding to the base sequence consisting of the 2917th to 3505th bases in the base sequence of SEQ ID NO: 1) of the clone Liver33 was cut out and inserted into the vector pGEX6P-1 (Amersham Pharmacia Biotech) for expressing a fusion protein with GST. E. coli BL21 was transformed with the plasmid to express a human SMG-1 fragment (corresponding to the amino acid sequence consisting of the 864th to 1059th amino acid residues in the amino acid sequence of SEQ ID NO: 2) as a fusion protein (molecular weight = approximately 50 kDa) with GST. This fusion protein produced in E. coli was also insoluble, and thus the immunogen was prepared in a manner similar to the case of preparing the immunogen of antiserum C3.

[0074]

To prepare antiserum N1 and antiserum N2, an approximately 0.7kbp of SmaI-HincII fragment (corresponding to the base sequence consisting of the 306th to 645th bases in the base sequence of SEQ ID NO: 1) derived from the clone AI005513 was inserted into the vector pGEX-6P (Amersham Pharmacia Biotech) for expressing a fusion protein with GST. The produced recombinant protein was purified from E. coli by the standard glutathione beads method, and was used as the immunogen.

In Fig. 4, the antigen sites are schematically shown. In Fig. 4, the regions (CR1 to CR6 in Fig. 2) with a high homology with C. elegans SMG-1 are shown by gray or black

squares. Further, in Fig. 4, "FRBH" means a sequence having homology with the FKBP12/rapamycin binding site (FKBP12/rapamycin binding homology), "PIKK" means a phosphatidyl inositol kinase (PIK) related kinase, and "PIKK-C" means a carboxyl terminal portion of the PIKK catalytic region. Further, the letters "N", "L", "C", and "P" mean the antigen sites used for preparing antisera N1 and N2, antisera L1 and L2, antiserum C3, and antiserum P1, respectively.

[0075]

Example 5: Detection of SMG-1 Protein in Various Animal Cells or Various Animal Tissues

(1) Detection of SMG-1 Protein in Various Animal Cell lysates by Western Blotting

HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 7% fetal bovine serum, and were ultrasonicated in a lysis buffer F [20 mmol/L Tris-HCl (pH7.5), 0.25 mmol/L sucrose, 1.2 mmol/L EGTA, 20 mmol/L β-mercapto ethanol, 1 mmol/L sodium orthovanadate, 1 mmol/L sodium pyrophosphate, 1 mmol/L sodium fluoride, 1% triton X-100, 0.5% nonidet P-40, 150 mmol/L NaCl, 1 mmol/L PMSF (phenylmethylsulfonyl fluoride), 10 μg/mL leupepsin, and 2 μg/mL aprotinin] to prepare a cell lysate.

[0076]

Similarly, various animal cell lysates were prepared for various cell lines derived from human, simian, mouse, and rat. More particularly, as the human cell lines, HeLa (ATCC: CCL-2), 293 (ATCC: CCL1573), HepG2 (ATCC: HB-8065), Jurkat [Schuneider, U. et al., Int. J. Cancer, 19, 621-626 (1977)], U937 [Sundstrom, C. et al., Int. J. Cancer, 17, 565 (1976)], HL-60 [Collins, S. J. et al., Nature, 270, 347 (1977)], and HPB-ALL [Morikawa, S. et al., Int. J. Cancer, 21, 166 (1978) were used. As the simian cell line, COS1 (ATCC: CRL1650) was used. As the mouse cell lines, NIH3T3 (ATCC: CRL1658), C3H10T1/2 (ATCC: CCL226), and C2C12 were used. As the rat cell lines, 3Y1 [Samdineyer, S. et al, Cancer Res., 41, 830 (1981)] and L6 [Yaffe, D. et al., Proc. Natl. Acad. Sci. USA, 61, 477-483 (1968)] were used.

[0077]

For the resulting various animal cell lysates (corresponding to 20 pg of protein), SDS-PAGE was performed at the gel concentrations of 5.5% and 12.5%, and then Western blotting was carried out using antiserum P1, antiserum C3, antiserum L1, antiserum L2, antiserum N1, and antiserum N2, and a preimmunized serum for control.

The results of use of antiserum P1, antiserum C3, antiserum L2, and antiserum N1 for the HeLa cell lysate are shown in Fig. 5. The results of use of antiserum P1 and antiserum C3 for various animal cell lysates are shown in Fig. 6.

In Fig. 5 and Fig. 6, "WB" means Western blotting. In Fig. 5, "pre" means the preimmunized serum. In Fig. 6, the arrow marks at the top in the "WB:C3" column or "WB:P1" column show p430, and the arrow marks at the bottom in the "WB:C3" column or "WB:P1" column show p400.

[0078]

In all antisera other than antiserum N1 and antiserum N2, two protein bands of 400 kDa and 430 kDa were antiserum-specifically detected. Hereinafter, the SMG-1 protein having the molecular weight of 400 kDa will be sometimes referred to as p400, and the SMG-1 protein having the molecular weight of 430 kDa will be sometimes referred to as p430. Further, in the two mouse cell lines NIH3T3 and C3H10T1/2, a protein band of 460 kDa was detected in addition to the two bands of 400 kDa and 430 kDa.

On the other hand, in the antiserum N1 and antiserum N2, only the 430 kDa band was detected. Therefore, the 400 kDa band is considered to be an SMG-1 molecule in which an N-terminal portion of human SMG-1 is deleted.

To prove this hypothesis, the nucleotide sequence of the hSMG-1 cDNA was carefully examined, whereupon the presence of the methionine (Met) codon satisfying the translation start criteria of Kozak at the 129th position became clear. The estimated ORF starting from the 129th Met is a 396,040 Da protein consisting of 3529 amino acids. Therefore, it is probably believed that p400 is a product of the ORF starting from the 129th second methionine.

[0079]

(2) Detection of SMG-1 Protein by Western Blotting in Cell Lysates Derived From Various Animal Tissues

With various tissues derived from rat and mouse, Western blotting was carried out using antiserum C3. Tissues were taken from animals by surgery, quickly frozen in liquid nitrogen, and powdered by crushing. Each powder was solubilized in a 1×SDS sample buffer, and then Western blotting was performed using 20 μg of protein from each tissue.

[0080]

The results are shown in Fig. 7. In Fig. 7, "WB" means Western blotting, the upper arrow mark indicates p430, and the lower arrow mark indicates p400. As the rat tissues, the heart, cerebrum, cerebellum, lung, liver, skeletal muscle, kidney, spleen, thymus, prostate, ovary, testis, and colon were used, and as the mouse tissue, the placenta was used.

In all tissues, two bands of the 400 kDa protein (p400) and the 430 kDa protein (p430) were detected. In the mouse placenta, a 460 kDa protein band was also detected in addition to the two 400 kDa and 430 kDa bands, but the 460 kDa band was a nonspecific signal.

[0081]

Example 6: Confirmation of Protein Kinase Activity of Human SMG-1 (Immunoprecipitate of Human HeLa Cell lysate by Anti-human SMG-1 Antiserum)

(1) Detection of SMG-1 Protein by Western Blotting in Immunoprecipitate of Human HeLa Cell Lysate by Various Human SMG-1 Antisera

The HeLa cell lysates obtained in a manner similar to that in the Example 5(1) were immunoprecipitated using antiserum N1, antiserum L2, and antiserum C3, and a preimmunized antiserum for control, respectively. The immunoprecipitation was performed by adding each antiserum to the cell lysate, allowing it to stand at 4°C for 2 hours to form an immunocomplex, adding protein A sepharose CL-4B (Amersham Pharmacia Biotech), allowing it to stand for a further 2 hours to bond the immunocomplex, and recovering the protein A sepharose CL-4B by centrifugation. For each

immunoprecipitate, SDS-PAGE was performed at a gel concentration of 5.5%, and Western blotting was performed using antiserum C3.

[0082]

The results are shown in Fig. 8. In Fig. 8, "WB" means Western blotting, and "^{32}P" means the results of autoradiography in Example 6(2). Further, "pre" means the preimmunization serum, and "IP" means the immunoprecipitate. Further, the arrow at the top side in the "^{32}P" column shows p430, and the arrow at the bottom side in the "^{32}P" column shows p400.

As shown by the "WB:C3" column of Fig. 8, while two protein bands of 400 kDa and 430 kDa were detected by the antiserum C3 from the immunoprecipitate of antiserum L2 or antiserum C3, only the protein band of 430 kDa was detected by the antiserum C3 from the immunoprecipitate of the antiserum N1.

[0083]

(2) Confirmation of Protein Kinase Activity of Immunoprecipitates of Human HeLa Cell Lysates by Various Human SMG-1 Antisera

The immunoprecipitates obtained in the Example 6(1) were washed with a lysis buffer F containing 0.25 mol/L LiCl, and then washed two times with a 1×kinase reaction buffer [10 mmol/L HEPES-KOH (pH7.5), 50 mmol/L β-glycerophosphoric acid, 50 mmol/L NaCl, 1 mmol/L dithiothreitol (DTT), and 10 mmol/L $MnCl_2$].

To each of the washed immunoprecipitates, 25 μL of 2×kinase reaction buffer (that is, two-fold concentrations of the above kinase reaction buffer) was added. The phosphorylation reaction was started by adding 10 mmol/L ATP and 370kBq [γ-^{32}P] ATP (6000 Ci/mmol; Amersham Pharmacia Biotech) in equal amounts (25 μL) and continued, with occasional stirring, at 30°C for 30 minutes. The final reaction amount was maintained at 50 μL, then 25 μL of a 4×SDS sample buffer was added to stop the reaction. SDS-PAGE was performed at gel concentrations of 5.5% and 12.5%, and then autoradiography was carried out to detect the phosphorylated proteins. The phosphorylation strength of

each protein was measured by an Image Analyzer BAS2000 (Fuji Film).

[0084]

The results are shown in Fig. 8. As shown in the "^{32}P" column of Fig. 8, in the immunoprecipitate by antiserum L2 or antiserum C3, phosphorylation proteins of the molecular weights 430 kDa and 400 kDa were detected. Proteins of the molecular weights 430 kDa and 400 kDa are believed to be human SMG-1, and thus it was found that human SMG-1 has an autophosphorylation activity.

[0085]

Example 7: Expression of Fusion Protein of Human SMG-1 Protein Fragment and One-Amino-Acid-Substituented Mutant

In this example, expression vectors were prepared for expressing (1) a fusion protein (hereinafter referred to as "6H-hSMG-1") of the human SMG-1 protein partial fragment having the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2, and the His tag consisting of the amino acid sequence of SEQ ID NO: 8 [including six continuous histidine (His) residues] and (2) a kinase-deficient mutant [hereinafter referred to as "6H-hSMG-1(DA)"] in which the asparatic acid (D) corresponding to the 2331st asparatic acid in the amino acid sequence of SEQ ID NO: 2 in the 6H-hSMG-1 is replaced with alanine (A).

[0086]

(1) Construction of Vector for Expression of Fusion Protein (6H-hSMG-1) of Human SMG-1 Protein Fragment and His Tag

An expression vector for expressing 6H-hSMG-1 was constructed by the following procedure.

The cDNA clone including a part (corresponding to the amino acid sequence consisting of the 107th to 3657th amino acids in the amino acid sequence of SEQ ID NO: 2) of the full-length of the hSMG-1 cDNA was digested by restriction enzymes HpaI and XhoI, and the 11kbp DNA fragment was purified. The DNA fragment was inserted into the SmaI/XhoI site of an expression vector SR6H [a modified SRD vector having a base sequence encoding the His tag upstream of the multicloning site (MCS)] to obtain a vector SR6H-hSMG-1 for

expressing the recombinant human SMG-1.

[0087]

(2) Construction of Vector for Expressing One-Amino-Acid-Substituented Mutant [6H-hSMG-1(DA)] of 6H-hSMG-1

Next, a vector SR6H-hSMG-1 (DA) for expressing 6H-hSMG-1 (DA) was obtained by using the above expression vector SR6H-hSMG-1 and a commercially available kit (Chameleon Mutagenesis Kit, Stratagen).

[0088]

(3) Confirmation of Expression of 6H-hSMG-1 and 6H-hSMG-1(DA) and Protein Kinase Activity in Vitro

After 293T cells were cultured using Dulbecco's modified Eagle's medium (DMEM; GibcoBRL), the cells were transfected with the expression vector SR6H-hSMG-1 prepared in Example 7(1) or the expression vector SR6H-hSMG-1(DA) prepared in Example 7(2). In this connection, as a control, transfection was also performed using the vector SR6H. After two days from the transfection, the cells were collected and lysed with the lysis buffer F.

Except for using an anti-polyhistidine antibody (His-Tag; Novagen), immunoprecipitation of each cell lysate was carried out in accordance with the procedure described in Example 6(1), and then the protein kinase activity in each of the resulting immunoprecipitates was measured in accordance with the procedure described in the Example 6(2). Further, Western blotting was also performed using the immunoprecipitates obtained by the immunoprecipitation.

[0089]

The results are shown in Fig. 9. In Fig. 9, "WB:anti-His" shows the results of Western blotting by the anti-polyhistidine antibody, and "^{32}P" shows the results of autoradiography. Further, "vector" means the results in the case of use of the vector SR6H (control), "hSMG-1 WT" means the results in the case of use of the vector SR6H-hSMG-1, and "hSMG-1 DA" means the results in the case of use of the vector SR6H-hSMG-1 (DA). Further, the arrow mark in the "^{32}P" column shows 6H-hSMG-1.

As shown in Fig. 9, both 6H-hSMG-1 and 6H-hSMG-1(DA) were immunoprecipitated by the anti-polyhistidine antibody.

Further, It was shown that the asparatic acid in the hSMG-1 corresponding to the 2331st asparatic acid in the amino acid sequence of SEQ ID NO: 2 (corresponding to the 2475th asparatic acid known to be essential for the kinase activity in ATR) is necessary for the kinase activity. As shown in Fig. 9, 6H-hSMG-1 obtained by the immunoprecipitation exhibits a mobility of approximately 400 kDa, and has a distinctive kinase activity. These results clearly show that 6H-hSMG-1 has a distinctive autophosphorylation activity.

[0090]

Example 8: Confirmation of Involvement of SMG-1 in PTC Dependent Degradation of β-globin mRNA

(1) Construction of Reporter Gene Plasmid

It was confirmed that, in C. elegans, seven types of smg genes are involved in NMD. The inventor made the unexpected discovery that a novel member of the PIKK family exhibits a similarity in overall sequence to C. elegans SMG-1, and thereby decided to investigate whether or not hSMG-1 is involved in the NMD of mammals. To this end, a reporter gene (Fig. 10) having a gene sequence with or without a PTC at the 39th codon of human β-globin (BGG) arranged downstream of the CMV promoter was constructed as follows. In this construction, the CMV promoter is under the control of the upstream tetracycline-responsive element (TRE) sequence. Further, when introduced into a cell line having a plasmid pTet OFF, the transcription from this reporter gene is stopped specifically and quickly in the presence of tetracycline or its derivative (doxycycline). In Fig. 10, an exon is shown by a square, and an intron is shown by a straight line.

[0091]

To prepare a reporter gene plasmid pTRE BGG WT (PTC is absent at the 39th codon of BGG), a human β-globin gene fragment was amplified from a human gene library (Clonetech) by PCR, and was inserted into a pTRE vector (Clonetech). Further, a nonsense mutation of the human β-globin gene at the codon 39 was induced by the standard procedure to produce a reporter gene plasmid pTRE BGG PTC (PTC is present

at the 39th codon of BGG).

[0092]

(2) Evaluation of Amount of Accumulation of Reporter mRNA by Northern Blotting

A cell line HeLa Tet-OFF (Clonetech) or a cell line MEF Tet-OFF (Clonetech) was transfected with the reporter plasmid BGG-WT or the reporter plasmid BGG-39PTC prepared in the Example 8(1) together with a CAT plasmid as the internal standard, and was incubated in the absence of doxycycline, and then the accumulation of the BGG mRNA was evaluated by Northern blotting.

More particularly, as a transfection reagent, polyfectin (QIAGEN) was used in the case of the cell line HeLa Tet-OFF, and effectin (QIAGEN) was used in the case of the cell line MEF Tet-OFF. After 24 hours from the transfection, cells were re-inoculated in six 10 cm dishes and cultured in the absence of doxycycline for further 24 hours. The transcription from the reporter was stopped by adding 50 ng/mL of doxycycline, the cells were collected at the periods of 0 hour, 0.5 hour, 1 hour, or 3 hours, and then each of the total RNA was isolated. The amounts of BGG mRNA and CAT mRNA from equal amounts (2 μg) of cells were evaluated by Northern blotting using a BGG probe and a CAT probe.

[0093]

The results are shown in Fig. 11. In Fig. 11, "WT" means the results of the case of using the reporter plasmid BGG-WT, and "39PTC" means the results of the case of use of the reporter plasmid BGG-39PTC. Further, "BG" means the results obtained by the BGG probe, and "CAT" means the results obtained by the CAT probe.

As shown in Fig. 11, in both cell lines, the accumulation of mRNA of BGG-WT (that is, BGG without PTC) was more abundant than the accumulation of BGG-39PTC (that is, BGG with PTC at the 39 position).

[0094]

(3) Confirmation of Effect of 6H-hSMG-1 and 6H-hSMG-1(DA) on Accumulation of Reporter mRNA

The procedure in Example 8(2) was repeated except for

transfecting either the expression vector SR6H-hSMG-1 prepared in the Example 7(1) or the expression vector SR6H-hSMG-1(DA) prepared in the Example 7(2) at the same time.

The results relating to BGG-39PTC in the HeLa Tet-OFF cells are shown in Fig. 12 and Fig. 13. In Fig. 12 and Fig. 13, "vector" or "vec" means the results in the case of use of the vector SR6H (control), "hSMG-1 WT" or "WT" means the results in the case of use of the vector SR6H-hSMG-1, and "hSMG-1 DA" or "DA" means the results in the case of use of the vector SR6H-hSMG-1 (DA). Further, "BG" means the results obtained by the BGG probe, and "CAT" means the results obtained by the CAT probe. Further "39PTC" means the results in the case of use of the reporter plasmid BGG-39PTC.

When 6H-hSMG-1 (DA) is overexpressed, the accumulation of the BGG-39PTC transcripts is amplified, while when 6H-hSMG-1 is overexpressed, the amount of stable state mRNA encoding BGG-39PTC is reduced, compared with introduction of the vector SR6H (control). These results provide powerful proof supporting the fact that hSMG-1 and its inherent protein kinase activity are involved in the PTC dependent decay of the BGG mRNA.

[0095]

Next, to further confirm this fact, the effects of overexpression of 6H-hSMG-1 or 6H-hSMG-1(DA) in the half life of mRNA of BGG WT or BGG-39PTC were tested. The transcription from each of the BGG reporters was stopped by adding doxycycline to the incubator, the cells were collected at the predetermined periods (0 hour, 0.5 hour, 1 hour, 1.5 hours, 2 hours, and 3 hours), and then each of the BGG mRNA was measured.

[0096]

The results are shown in Fig. 14 to Fig. 17. In Fig. 14 to Fig. 17, "BGG WT" means the results in the case of use of the reporter plasmid BGG-WT, and "BGG PTC" means the results in the case of use of the reporter plasmid BGG-39PTC. Further, "vector" or "vec" means the results in the case of use of the vector SR6H (control), "hSMG-1 WT" or "WT" means the results in the case of use of the vector SR6H-hSMG-1,

and "hSMG-1 DA" or "DA" means the results in the case of use of the vector SR6H-hSMG-1(DA). Further, "Dox." means doxycycline, "BG" means BGG, and "18S" means 18S libosome RNA.

The half life of BGG WT appears to be extremely long, as already reported [Sun, X. et al., Proc. Natl. Acad. Sci. USA, 95, 10009-10014 (1998)], and further is not affected by the expression of either 6H-hSMG-1 or 6H-hSMG-1(DA). On the other hand, the half life of BGG-39PTC is greatly shortened by the overexpression of 6H-hSMG-1 and becomes longer due to the overexpression of 6H-hSMG-1(DA). When combining these results with the above results, it is clearly shown that 6H-hSMG-1 is involved in the decay of PTC-dependent BGG mRNA. Further, these results also show that the kinase activity of 6H-hSMG-1 plays an important role in the NMD of mammals.

[0097]

Example 9: Phosphorylation of hUPF1/SMG-2 by 6H-hSMG-1 in vitro

An experiment by Perlick [Perlick, H. A. et al., Proc. Natl. Acad. Sci. USA, 93, 10928-10932 (1996)] identified hUpf1 (a human homolog of yeast Upf1). Further, using a point mutation of the helicase domain of hUpf1, Sun et al. showed that hUpf1 is involved in the NMD of mammals [Sun, X. et al., Proc. Natl. Acad. Sci. USA, 95, 10009-10014 (1998)]. More recently, Anderson confirmed that C. elegans SMG-2 protein is a homolog of Upf1 in C. elegans [Page et al., Mol. Cell. Biol., 19, 5943-5951 (1999)]. SMG-2 is a phosphorylated protein. Further, of extreme importance, another six types of smg genes can be classified into two groups based on the effects of mutation in the phosphorylated state of SMG-2. In the mutants of smg-1, smg-2, and smg-3, SMG-2 in the phosphorylated state was not detected. In the mutants of smg-5, smg-6, and smg-7, phosphorylated SMG-2 was accumulated at a high level.

[0098]

(1) Confirmation of Phosphorylation of Full-length hUpf1/SMG-2 Fusion Protein by 6H-hSMG-1

To test the possibility that hSMG-1 directly phosphorylates hUpf1/SMG-2, the HA tagged hUpf1/SMG-2

(hereinafter referred to as HA-hUpf1/SMG-2) was expressed in 293T cells, and HA-hUpf1/SMG-2 was purified.

More particularly, first, an expression vector for expressing HA-hUpf1/SMG-2 was prepared by the following procedure. That is, an SR vector [Hirai, S. et al., Oncogene, 12, 641-650 (1996)] was modified by inserting the HA tag at the multicloning site (MCS) and upstream thereof to obtain a vector SRHAI. Into the MCS of the obtained vector SRHAI, cDNA encoding the full-length of hUpf1/SMG-2 was inserted to obtain an expression vector SRHAI-hUpf1/SMG-2. More particularly, the vector SRHAI was cleaved by restriction enzyme BglII, and then blunted. Into the blunted vector, the cDNA clone KIAA0221, which had been cleaved by restriction enzymes XhoI and BlpI and then blunted, was inserted.

[0099]

Then, 293T cells were transfected with the obtained expression vector SRHAI-hUpf1/SMG-2. Two days after the transfection, the cells were collected and lysed in the lysis buffer F. Anti-HA affinity beads (Rosche) were added to the lysate. After one hour, the beads were washed with the lysis buffer F three times and washed with a washing buffer [20 mmol/L Tris-HCl (pH7.5), 0.1 mol/L NaCl, 0.1 mmol/L EDTA, and 0.05% Tween20] three times. The resulting washed beads were treated in the washing buffer containing 1 mg/mL HA peptide (YPYDVPDYA) at 37°C to elute the binding protein. Next, dialysis in 1×PBS containing 10% glycerol and 1 mmol/L DTT was carried out to obtain HA-hUpf1/SMG-2.

[0100]

On the other hand, 6H-hSMG-1 and 6H-hSMG-1 (DA) were purified from cDNA-transfected 293T cells transfected by the expression vector SR6H-hSMG-1 prepared in Example 7(1) or the expression vector SR6H-hSMG-1 (DA) prepared in Example 7(2) in accordance with the procedure described in Example 7(3).

The phosphorylation reaction was performed in accordance with the procedure described in Example 6(2), except for adding HA-hUpf1/SMG-2 prepared in Example 9(1) to the 2×kinase reaction buffer as a substrate.

[0101]

The results are shown in Fig. 18. In Fig. 18, "vector" means the results in the case of use of the vector SR6H (control), "hSMG-1 WT" means the results in the case of use of the vector SR6H-hSMG-1, and "hSMG-1 DA" means the results in the case of use of the vector SR6H-hSMG-1 (DA). "anti-His" means the results of Western blotting by the anti-polyhistidine antibody, "^{32}P" means the results of autoradiography, and "CBB" means the results obtained by the Coomassie Brilliant Blue (CBB) staining.

As shown in Fig. 18, purified 6H-hSMG-1 phosphorylated HA-hUpf1/SMG-2. This suggests that, at least in the system using the purified substance, hUpf1/SMG-2 becomes a direct substrate of hSMG-1. Kinases belonging to the PIKK family phosphorylate the serine or threonine residue in the SQ or TQ motif [Kim, S. T. et al., J. Biol. Chem., 274, 37538-37543 (1999)]. Of interest, hUpf1/SMG-2 contains a repetition of the SQ motif in the C-terminal region [Page et al., Mol. Cell. Biol., 19, 5943-5951 (1999)]. Taking into consideration the fact that hSMG-1 encodes the kinase belonging to the PIKK family, this suggests that the SQ motif is the target of hSMG-1.

[0102]

(2) Confirmation of Phosphorylation by 6H-hSMG-1 in Fusion Protein of hUpf1/SMG-2 Partial Fragment (1)

To confirm the above hypothesis, a series of maltose binding protein (MBP) fusion proteins containing the fragmentated hUpf1/SMG-2 was constructed and purified.

More particularly, three types of cDNA fragments cut from SRHAI-hUpf1/SMG-2 [prepared in Example 9(1)] containing cDNA encoding hUpf1/SMG-2, that is, a cDNA fragment (1.4kbp, BgIII-Eco47III fragment, corresponding to the amino acid sequence consisting of the 1st to 462nd amino acids of hUpf1/SMG-2) encoding a partial fragment at the N-terminal side, a cDNA fragment (1.0kbp, Eco47IH-Eco47II fragment, corresponding to the amino acid sequence consisting of the 463rd to 800th amino acids of hUpf1/SMG-2) encoding a partial fragment in the intermediate region, and a cDNA fragment (1.4kbp, Eco47llI-BstZ17I fragment, corresponding

to the amino acid sequence consisting of the 801st to 1118th amino acids of hUpf1/SMG-2) encoding a partial fragment at the C-terminal side, were inserted into a pMaI-c2 vector (New England Biolabs) to obtain the expression vectors pMBP-hSMG-2 N, pMBP-hSMG-2 M, and pMBP-hSMG-2 C, respectively.

[0103]

The obtained MBP fusion proteins were all extremely insoluble in E. coli, and thus the recombinant proteins were purified from inclusion bodies as follows. That is, the collected cells were suspended in an ultrasonication buffer [50 mmol/L TrisHCl (pH8.0), 50 mmol/L NaCl, 1 mmol/L EDTA, 1 mmol/L DTT, and 1% triton X-100] containing 2 μg/mL aprotinin, 10 μg/mL leupeptin, 2 mmol/L PMSF, and 50 mmol/L benzamidine, and were ultrasonicated. Each precipitate (mostly inclusion bodies) obtained by centrifugation at 10000×g was washed with a washing solution (0.5% triton X-100 and 1 mmol/L EDTA) five times. The washed precipitate was suspended in a denaturation buffer [8 mol/L urea, 50 mmol/L TrisHCl (pH8.0), 1 mmol/L DTT, and 1 mmol/L EDTA], and allowed to stand at room temperature for 1 hour. The supernatant obtained by centrifugation at 10000×g was dialyzed for 1 hour in a denaturation buffer containing 4 mol/L urea, then was dialyzed for 1 hour in a denaturation buffer containing 2 mol/L urea, and further was dialyzed overnight in the ultrasonication buffer. MBP fusion proteins (i.e., the fusion proteins of the partial fragment of Upf1/SMG-2 at the N-terminal side, the partial fragment in the intermediate region, or the partial fragment at the C-terminal side, with MBP) renaturated by this treatment was recovered and purified using an amylose resin (New England Biolabs) in accordance with the attached manual.

[0104]

The phosphorylation reaction was performed in accordance with the procedure described in Example 6(2), except for adding as a substrate each MBP fusion protein to the 2×kinase reaction buffer and using, as hSMG-1, 6H-hSMG-1 prepared in accordance with the procedure described in Example 7(3).

The results are shown in Fig. 19 and Fig. 20. In Fig.

20, "CBB" means the results by CBB staining, while "^{32}P" means the results of autoradiography. Further, the numerals shown under the autoradiograms are relative values when using the intensity of the autoradiogram in the fusion protein of pMBP-hSMG-2 C and MBP as 100.

As shown in Fig. 20, the fragments of hUpf1/SMG-2 at the C-terminal side and at the N-terminal side performed the role of good substrates for hSMG-1. The results of the fragment of hUpf1/SMG-2 at the C-terminal side being phosphorylated, taking into consideration the Page et al. report (that is, hUpf1/SMG-2 contains a repetition of the SQ motif at the C-terminal region), lead to the prediction that the SQ motif is phosphorylated. Further, as a result of the fragment of hUpf1/SMG-2 at the N-terminal side being phosphorylated, it is believed that there are plural SQ motifs at the N-terminal region and that there is a possibility that these sites are phosphorylated.

[0105]

(3) Confirmation of Phosphorylation by 6H-hSMG-1 in Fusion Protein of hUpf1/SMG-2 Partial Fragment (2)

Next, to further clarify the above point, another series of GST fusion proteins was prepared. In this example, fusion proteins in which 14mer peptides consisting of the SQ or TQ deduced motifs in hUpf1/SMG-2 and the surrounding 12 amino acid residues were fused downstream of GST were prepared.

More particularly, each DNA encoding a 14mer peptide containing T28 (that is, the 28th threonine in hUpf1/SMG-2), T325 (that is, the 325th threonine), S474 (that is, the 474th serine), S681 (that is, the 681st serine), S1078 (that is, the 1078th serine), or S1096 (that is, the 1096th serine), or DNA encoding the 14mer peptide (control) containing S15 in the p53 protein (the 15th serine in the p53 protein) was inserted into a vector pGEX 6P (Amersham Pharmacia Biotech) to prepare each expression vector. Each GST fusion protein was purified from E. coli transformed with each expression vector by the standard glutathione beads method.

[0106]

The amino acid sequences of the 14mer peptides are shown in Fig. 21. In Fig. 21, "T28" means the amino acid sequence of the 14mer peptide part in the fusion protein of GST and the 14mer peptide containing T28. Similarly, "T325", "S474", "S681", "S1078", and "S1096" mean the amino acid sequences of the 14mer peptide parts in the fusion proteins of GST and the 14mer peptides containing T325, S474, S681, S1078, and S1096, respectively. "p53 S15" means the amino acid sequence of the 14mer peptide part in the fusion protein of GST and the 14mer peptide (control) containing S15.

[0107]

The phosphorylation reaction was performed in accordance with the procedure described in the Example 6(2), except for adding as the substrate each GST fusion protein to the 2×kinase reaction buffer and using, as hSMG-1, 6H-hSMG-1 prepared in accordance with the procedure described in Example 7(3).

The results are shown in Fig. 22. In Fig. 22, "T28" means a fusion protein of the 14mer peptide including T28 and GST. Similarly, "T325", "S474", "S681", "S1078", and "S1096" mean fusion proteins of the 14mer peptides including T325, S474, S681, S1078, and S1096, and GST, and "p53 S15" means a fusion protein of the 14mer peptide (control) including S15 in the p53 protein and GST. "S1078A" means a point mutant in which the 1078th serine in "S1078" is replaced with alanine. Further, "CBB" means the results of CBB staining, while "^{32}P" means the results of autoradiography. Further, the numerals shown at the bottom of the autoradiograms are relative values in the case of using the strength of the autoradiogram in the fusion protein (p53 S15) of 14mer peptide including S15 in the p53 protein and GST as 100.

[0108]

As shown in Fig. 22, the control construct encoding the SQ motif in the p53 protein was phosphorylated by hSMG-1. Further, the GST fusion protein including S1078 or the GST fusion protein including S1096 [hereinafter referred to as an hUpf1/SMG-2 fusion protein (S1096)] was efficiently

phosphorylated by 6H-hSMG-1. These results establish that 6H-hSMG-1 phosphorylates the serine residues in S1078 and S1096 as the SQ motifs of hUpf1/SMG-2, at least in vitro.

[0109]

Example 10: Confirmation of Phosphorylation of hUpf1/SMG-2 by SMG-1 in Cells

Considering the results obtained in the Example 9 (that is, the result that 6H-hSMG-1 phosphorylates hUpf1/SMG-2 in vitro) together with the results in the C. elegans smg genes, an interesting possibility is raised that hSMG-1 phosphorylates hUpf1/SMG-2 even in vivo and further, that the phosphorylation plays a fundamental role in NMD. As a first step for evaluating this possibility, the phosphorylation of hUpf1/SMG-2 was tested in vivo.

[0110]

The HeLa cells were treated with various concentrations of okadaic acid (OA; Calbiochem) for 4.5 hours, and then were recovered and dissolved in the 1×SDS sample buffer. After 6% SDS-PAGE was performed, Western blotting using an anti-hUpf1/SMG-2 antibody was performed to determine the mobility shift of hUpf1/SMG-2.

The results are shown in Fig. 23. When HeLa cells are treated with okadaic acid (OA), a phosphatase inhibitor, as a result, an upwardly shifted band of hUpf1/SMG-2 appears. In Fig. 23, the position of the shifted band is marked by an asterisk. Further, the "anti-hUPF1/SMG-2" in Fig. 23 means the results obtained by Western blotting using the anti-hUpf1/SMG-2 antibody.

[0111]

To show that the upward shift of hUpf1/SMG-2 induced by OA arises due to phosphorylation, the immunopurified hUpf1/SMG-2 was treated with alkaline phosphatase, then the mobility in SDS-PAGE was tested as follows.

That is, HeLa cells treated for 4.5 hours in the presence or absence (that is, only the medium) of 50 nmol/L okadaic acid were recovered, lysed in the lysis buffer F containing 1 μmol/L mycrocystin LR (Calbiochem) and 10 nmol/L okadaic acid, and then immunoprecipitated using an anti-hUpf1/SMG-2 serum. The reason why the mycrocystin and

okadaic acid were added to the lysis buffer F was to prevent the once phosphorylated protein from being dephosphorylated during immunoprecipitation.

The immunoprecipitate was washed in the lysis buffer F and a dephosphorylation buffer [50 mmol/L Tris-HCl (pH9.0) and 1 mmol/L $MgCl_2$], and then suspended in 50 μL of the dephosphorylation buffer. Calf intestine alkaline phosphatase (CIAP; Takara Shuzo) was added in an amount of 0 unit (that is, not added) or 60 units to start the reaction. The mixture was incubated at 37°C for 1 hour, then the SDS sample buffer was added to stop the reaction. After 6% SDS-PAGE was performed, the mobility shift of hUpf1/SMG-2 was determined by Western blotting using the anti-Upf1/SMG-2 antibody.

[0112]

The results are shown in Fig. 24. In Fig. 24, "OA" means the results in the case of using the immunoprecipitate derived from cells treated with okadaic acid, while "medium" means the results in the case of using the immunoprecipitate derived from cells in the absence of okadaic acid. Further, "anti-hUPF1/SMG-2" means the results obtained by Western blotting using the anti-hUpf1/SMG-2 antibody. Further, "hUPF1-P" means phosphorylated hUpf1/SMG-2, while "hUPF1" means unphosphorylated hUpf1/SMG-2.

The upwardly shifted band disappeared in the case of treating the immunoprecipitate by phosphatase (CIAP). This shows that the upward shift of hUpf1/SMG-2 occurring due to the OA treatment is phosphorylation.

[0113]

Next, to analyze the overexpressed hUpf1/SMG-2, 293T cells were transfected by the expression vector SRHAI-hUpf1/SMG-2 for expressing HA-hUpf1/SMG-2 prepared in Example 9(1) and the expression vector SR6H-hSMG-1 or vector SR6H-hSMG-1 (DA) prepared in Example 7(1). The cells were cultured for 4 hours in the presence or absence of 50 nmol/L okadaic acid. The cells were recovered and then dissolved in the 1×SDS sample buffer. The mobility shift of hUpf1/SMG-2 was determined by the Western blotting using an anti-HA antibody (12CA5; Boehringer).

[0114]

The results are shown in Fig. 25. In Fig. 25, "vector" means the results when using the vector SR6H (control), "hSMG-1 WT" means the results when using the vector SR6H-hSMG-1, and "hSMG-1 DA" means the results when using the vector SR6H-hSMG-1 (DA). Further, "anti-His" means the results of Western blotting using the anti-polyhistidine antibody. Further, "HA hUPF1-P" means phosphorylated HA-hUpf1/SMG-2, while "HA hUPF1" means unphosphorylated HA-hUpf1/SMG-2. In Fig. 25, the position of the shifted HA-hUpf1/SMG-2 is marked by an asterisk.

In a manner similar to the case of only the vector SR6H (control), when overexpressing 6H-hSMG-1 (DA), no OA-induced upward shift of the exogenous HA tagged hUpf1/SMG-2 was observed. However, when 6H-hSMG-1 was overexpressed, the OA-induced upward shift of the HA tagged hUpf1/SMG-2 was greatly amplified.

[0115]

Example 11: Identification of Inhibitor Using 6H-hSMG-1 Protein Kinase Activity as Indicator

From past research into the PIKK family, inhibitors acting in this family of kinases are identified. As the identified inhibitors, for example, wortmannin [Sarkaria, S. N. et al., Cancer Res., 58, 4375-4382 (1998)] and caffeine [Sarkaria, S. N. et al., Cancer Res., 59, 4375-4382 (1999)] may be mentioned. Next, to evaluate the role of hSMG-1 in NMD in mammals and to evaluate the potential strategy of specific inhibition of NMD by pharmacological operations on cell, hUpf1/SMG-2 fusion protein (S1096) prepared in Example 9(3) [that is, fusion protein in which the 14mer peptide including the 1096th serine (S1096) is fused downstream of GST] was used as the endogenous substrate, to evaluate the effects of these inhibitors in the hSMG-1 kinase activity.

More particularly, 6H-hSMG-1 was prepared in accordance with the procedure described in Example 7(3). In the presence of various concentrations of wortmannin or caffeine shown in Fig. 26 and Fig. 27, the hUpf1/SMG-2 fusion protein (S1096) prepared in Example 9(3) was used as the substrate, to perform an in vitro kinase assay. That is, the

phosphorylation was performed in accordance with the procedure described in Example 6(2), except for adding the hUpf1/SMG-2 fusion protein (S1096) and wortmannin or caffeine to the 2×kinase reaction buffer and using, as hSMG-1, 6H-hSMG-1 prepared in accordance with the procedure described in Example 7(3).

[0116]

The results in the case of useing wortmannin are shown in Fig. 26, while the results in the case of useing caffeine are shown in Fig. 27. As shown in Fig. 26 and Fig. 27, both wortmannin and caffeine inhibited the kinase activity of 6H-hSMG-1 by IC50 values of approximately 60 nmol/L and 0.3 mmol/L, respectively. On the other hand, rapamycin did not inhibit hSMG-1 in the presence of purified recombinant FKBP12 (data not shown).

[0117]

Example 12: Confirmation of SMG-1 Inhibitor Inhibiting Phosphorylation of hUpf1/SMG-2 in Cells

Further, the effects of the two types of hSMG-1inhibitor can also be tested in the phosphorylation of endogenous hUpf1/SMG-2 in HeLa cells.

HeLa cells were pretreated for 30 minutes in the presence or absence of various concentrations of wortmannin, caffeine, or rapamycin shown in Fig. 28. Next, the cells were treated for 4.5 hours in the presence of wortmannin, caffeine, or rapamycin and in the presence or absence of 50 nmol/L okadaic acid. Cell lysates were prepared and analyzed by Western blotting using the anti-Upf1/SMG-2 antibody.

The results are shown in Fig. 28. In Fig. 28, "anti-hUPF1/SMG-2" means the results obtained from Western blotting using the anti-hUpf1/SMG-2 antibody. Further, "cont.", "wort.", "caff.", and "rap." show the results of a control (that is, in the absence of wortmannin, caffeine, and rapamycin), the results in the presence of wortmannin, the results in the presence of caffeine, and the results in the presence of rapamycin, respectively. Further, "hUPF1-P" means phosphorylated hUpf1/SMG-2, while "hUPF1" means unphosphorylated hUpf1/SMG-2.

As shown in Fig. 28, wortmannin and caffeine both inhibited the upward shift of hUpf1/SMG-2 in HeLa cells, while rapamycin did not. This result matches with the results in the purified system (that is, the results of Example 11).

[0118]

Example 13: Stabilization of Endogenous PTC mRNA by SMG-1 Inhibitor

(1) Stabilization of BGG Gene Product Containing Endogenous PTC by SMG-1 Inhibitor

If hSMG-1 plays an important role in the NMD of mammals, these hSMG-1 inhibitors should inhibit NMD. To test this, first, the reporter BGG systems utilizing the reporter plasmid BGG-WT or the reporter plasmid BGG-39 PTC prepared in Example 8(1) were applied.

More particularly, MEF-Tet OFF cells were transfected with the reporter plasmid BGG-WT or the reporter plasmid BGG-39 PTC, and re-inoculated in eight dishes. The cells were then treated for 4.5 hours in the presence of 50 ng/ml doxycycline by various concentrations of caffeine (caff.), wortmannin (wort.), rapamycin (rap.), or cyclohexamide (CHX) shown in Fig. 29.

[0119]

The Total RNA was analyzed by Northern blotting using the BGG probe. The results are shown in Fig. 29. In Fig. 29, "BG WT" means the results in the case of use of the reporter plasmid BGG-WT, "BG PTC" means the results in the case of use of the reporter plasmid BGG-39PTC, and "GAPDH" means the results in the case of use of the cDNA of glyceryl aldehyde-3-phosphate dehydrogenase as a probe. Further, "cont.", "caff.", "wort.", "rap.", and "CHX" show the results of the control (that is, in the absence of wortmannin, caffeine, rapamycin, and cyclohexamide), the results in the presence of caffeine, the results in the presence of wortmannin, the results in the presence of rapamycin, and the results in the presence of cyclohexamide, respectively.

As shown in Fig. 29, a protein synthesis inhibitor, CHX inhibited NMD. Further, BGG-39PTC mRNA (not BGG WT) was

accumulated. This result matches the observations as described above. Of importance, the hSMG-1 inhibitors, that is, caffeine and wortmannin, resulted in the accumulation of BGG 39PTC. From this result, pharmacological proof supporting the assertion that hSMG-1 is involved in the NMD of mammals was obtained.

[0120]

(2) Stabilization of Endogenous PTC p53 Gene Product by SMG-1 Inhibitor

NMD rescues cells from the accumulation of potentially toxic proteins produced from PTC mRNA, but NMD often eliminates mRNAs encoding fragmentated proteins with residual activity capable of partially rescuing an impaired phenotype caused due to the mutation. Therefore, at least in the cases of several PTC mutations, it is possible to provide a novel method of treatment for rescuing the genetic disorders, by specifically inhibiting NMD.

Next, as a first step for evaluating the possibilities of the method, the ability of the hSMG-1 inhibitors to specifically rescue the synthesis of fragmentated proteins was tested. As a model of a system for evaluating the possibility, the p53 gene was selected because cell lines having the mutation can be obtained. Two types of cell lines having PTCs, that is, Calu6 (lung adenocarcinoma cell line) including the PTC at the 196th codon and N417 (small cell lung adenocarcinoma cell line) including the PTC at the 1298th codon [Lehman TA, Cancer Research, 51, 4090-4096 (1991); Bodner SM, Oncogene, 7, 743-749 (1992)] were selected. The structure of the p53 gene and the PTC mutations of the cell lines Calu6 and N417 are schematically shown in Fig. 30. In Fig. 30, an exon is shown by a square.

[0121]

The Calu6 and N417 cells, and the A549 cells [lung adenocarcinoma cell line; Lehman TA, cancer research, 51, 4090-4096 (1991)] as the control were treated in the presence or absence of 2 μmol/L wortmannin (wort.) or 50 μg/mL cyclohexamide (CHX) (cont.) for 4.5 hours, and then were recovered. The prepared cell lysates and total RNAs were analyzed by Northern blotting using a p53 probe and

Western blotting using an anti-p53 antibody (DO-1; Calbiochem). A CBB image showing actin staining is also displayed.

[0122]

The results in the N417 and A549 cells are shown in Fig. 31. In Fig. 31, "cont.", "wort.", and "CHX" show the results of the control, the results in the presence of wortmannin, and the results in the presence of cyclohexamide, respectively.

As a result of treatment of N417 cells by wortmannin, the p53 298PTC mRNA and the fragmentated p53 protein both increased, but in the control A549 cells, neither the mRNA nor the protein increased.

[0123]

Further, the results in the case of treatment for 4.5 hours by various concentrations of wortmannin, cyclohexamide, or caffeine are shown in Fig. 32. In Fig. 32, "CHX" shows the results in the presence of cyclohexamide. The increase in the fragmentated p53 was also observed in the case of treatment of calu6 cells by an increased amount of wortmannin.

[0124]

[Effects of the Invention]

According to the polypeptide of the present invention, a convenient screening system for agents of treating and/or preventing a disease caused by one or more PTCs generated by a nonsense mutation can be provided. Further, the polynucleotide, expression vector, cell, and antibody of the present invention are useful in manufacturing the polypeptide of the present invention.

[0125]

[FREE TEXT IN SEQUENCE LISTING]

Features of "Artificial Sequence" are described in the numeric identifier <223> in the Sequence Listing. More particularly, the base sequence of SEQ ID NO: 8 in the Sequence Listing is a His tag containing six histidine residues.

[0126]

[Sequencing List]

<110> Ohno, Shigeo

<120> Novel SMG-1

<130> YLS01001P

<160> 8

<210> 1
<211> 13110
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (328)..(11301)

<400> 1
ggggaagcag tggccgtgtg agcgtgagga gctgccgcca ccgcctgctc ctcgtcctcc 60

tcgtcctccg gggccccagc gtcgtgggcc gcgcacggcc ctggaagaga cgtcgcctcg 120

ccttcatccg cctctctcac cgcgccgctc cctcgtcctg ccctgcgggc tcaggcggaa 180

cccggaacgg ccgtcctctt cccccgccct ccgccgccgc ctcctcctcc tccttctcgg 240

cttcctcctc agccccgggc cggagcgggg tgtcggcggc ggccggttcg ggcggcggcg 300

cttggccatg tcgtgtcggg gaaggta atg agc cgc aga gcc ccg ggg tct cgg 354

Met Ser Arg Arg Ala Pro Gly Ser Arg
1 5

ctg agc agc ggc ggc acc aac tat tcg cgg agc tgg aat gac tgg caa 402
Leu Ser Ser Gly Gly Thr Asn Tyr Ser Arg Ser Trp Asn Asp Trp Gln
10 15 20 25

ccc aga act gat agt gca tca gct gac cca ggt aat tta aaa tat tct 450
Pro Arg Thr Asp Ser Ala Ser Ala Asp Pro Gly Asn Leu Lys Tyr Ser
30 35 40

tca tcc aga gat aga ggt ggt tct tcc tct tac gga ctg caa cct tca 498
Ser Ser Arg Asp Arg Gly Gly Ser Ser Ser Tyr Gly Leu Gln Pro Ser
45 50 55

aat tca gct gtg gtg tct cgg caa agg cac gat gat acc aga gtc cac 546
Asn Ser Ala Val Val Ser Arg Gln Arg His Asp Asp Thr Arg Val His
60 65 70

gct gac ata cag aat gac gaa aag ggt ggc tac agt gtc aat gga gga 594
Ala Asp Ile Gln Asn Asp Glu Lys Gly Gly Tyr Ser Val Asn Gly Gly
75 80 85

tct ggg gaa aat act tat ggt cgg aag tcg ttg ggg caa gag ctg agg 642
Ser Gly Glu Asn Thr Tyr Gly Arg Lys Ser Leu Gly Gln Glu Leu Arg
90 95 100 105

gtt aac aat gtg acc agc cct gag ttc acc agt gtt cag cat ggc agt 690
Val Asn Asn Val Thr Ser Pro Glu Phe Thr Ser Val Gln His Gly Ser

110 115 120

cgt gct tta gcc acc aaa gac atg agg aaa tca cag gag aga tcg atg 738
Arg Ala Leu Ala Thr Lys Asp Met Arg Lys Ser Gln Glu Arg Ser Met
125 130 135

tct tat tct gat gag tct cga ctg tcg aat ctt ctt cgg agg atc acc 786
Ser Tyr Ser Asp Glu Ser Arg Leu Ser Asn Leu Leu Arg Arg Ile Thr
140 145 150

cgg gaa gac gac aga gac cga aga ttg gct act gta aag cag ttg aaa 834
Arg Glu Asp Asp Arg Asp Arg Arg Leu Ala Thr Val Lys Gln Leu Lys
155 160 165

gaa ttt att cag caa cca gaa aat aag ctg gta cta gtt aaa caa ttg 882
Glu Phe Ile Gln Gln Pro Glu Asn Lys Leu Val Leu Val Lys Gln Leu
170 175 180 185

gat aat atc ttg gct gct gta cat gac gtg ctt aat gaa agt agc aaa 930
Asp Asn Ile Leu Ala Ala Val His Asp Val Leu Asn Glu Ser Ser Lys
190 195 200

ttg ctt cag gag ttg aga cag gag gga gct tgc tgt ctt ggc ctt ctt 978
Leu Leu Gln Glu Leu Arg Gln Glu Gly Ala Cys Cys Leu Gly Leu Leu
205 210 215

tgt gct tct ctg agc tat gag gct gag aag atc ttc aag tgg att ttt 1026
Cys Ala Ser Leu Ser Tyr Glu Ala Glu Lys Ile Phe Lys Trp Ile Phe
220 225 230

agc aaa ttt agc tca tct gca aaa gat gaa gtt aaa ctc ctc tac tta 1074
Ser Lys Phe Ser Ser Ser Ala Lys Asp Glu Val Lys Leu Leu Tyr Leu
235 240 245

tgt gcc acc tac aaa gca cta gag act gta gga gaa aag aaa gcc ttt 1122
Cys Ala Thr Tyr Lys Ala Leu Glu Thr Val Gly Glu Lys Lys Ala Phe
250 255 260 265

tca tct gta atg cag ctt gta atg acc agc ctg cag tct att ctt gaa 1170
Ser Ser Val Met Gln Leu Val Met Thr Ser Leu Gln Ser Ile Leu Glu
270 275 280

aat gtg gat aca cca gaa ttg ctt tgt aaa tgt gtt aag tgc att ctt 1218
Asn Val Asp Thr Pro Glu Leu Leu Cys Lys Cys Val Lys Cys Ile Leu
285 290 295

ttg gtg gct cga tgt tac cct cat att ttc agc act aat ttt agg gat 1266
Leu Val Ala Arg Cys Tyr Pro His Ile Phe Ser Thr Asn Phe Arg Asp
300 305 310

aca gtt gat ata tta gtt gga tgg cat ata gat cat act cag aaa cct 1314
Thr Val Asp Ile Leu Val Gly Trp His Ile Asp His Thr Gln Lys Pro
315 320 325

tcg ctc acg cag cag gta tct ggg tgg ttg cag agt ttg gag cca ttt 1362
Ser Leu Thr Gln Gln Val Ser Gly Trp Leu Gln Ser Leu Glu Pro Phe
330 335 340 345

```
tgg gta gct gat ctt gca ttt tct act act ctt ctt ggt cag ttt ctg   1410
Trp Val Ala Asp Leu Ala Phe Ser Thr Thr Leu Leu Gly Gln Phe Leu
                350                 355                 360

gaa gac atg gaa gca tat gct gag gac ctc agc cat gtg gcc tct ggg   1458
Glu Asp Met Glu Ala Tyr Ala Glu Asp Leu Ser His Val Ala Ser Gly
            365                 370                 375

gaa tca gtg gat gaa gat gtc cct cct cca tca gtg tca tta cca aag   1506
Glu Ser Val Asp Glu Asp Val Pro Pro Pro Ser Val Ser Leu Pro Lys
        380                 385                 390

ctg gct gca ctt ctc cgg gta ttt agt act gtg gtg agg agc att ggg   1554
Leu Ala Ala Leu Leu Arg Val Phe Ser Thr Val Val Arg Ser Ile Gly
    395                 400                 405

gaa cgc ttc agc cca att cgg ggt cct cca att act gag gca tat gta   1602
Glu Arg Phe Ser Pro Ile Arg Gly Pro Pro Ile Thr Glu Ala Tyr Val
410                 415                 420                 425

aca gat gtt ctg tac aga gta atg aga tgt gtg acg gct gca aac cag   1650
Thr Asp Val Leu Tyr Arg Val Met Arg Cys Val Thr Ala Ala Asn Gln
                430                 435                 440

gtg ttt ttt tct gag gct gtg ttg aca gct gct aat gag tgt gtt ggt   1698
Val Phe Phe Ser Glu Ala Val Leu Thr Ala Ala Asn Glu Cys Val Gly
            445                 450                 455

gtt ttg ctc ggc agc ttg gat cct agc atg act ata cat tgt gac atg   1746
```

Val Leu Leu Gly Ser Leu Asp Pro Ser Met Thr Ile His Cys Asp Met
460 465 470

gtc att aca tat gga tta gac caa ctg gag aat tgc cag act tgt ggt 1794
Val Ile Thr Tyr Gly Leu Asp Gln Leu Glu Asn Cys Gln Thr Cys Gly
475 480 485

acc gat tat atc atc tca gtc ttg aat tta ctc acg ctg att gtt gaa 1842
Thr Asp Tyr Ile Ile Ser Val Leu Asn Leu Leu Thr Leu Ile Val Glu
490 495 500 505

cag ata aat acg aaa ctg cca tca tca ttt gta gaa aaa ctg ttt ata 1890
Gln Ile Asn Thr Lys Leu Pro Ser Ser Phe Val Glu Lys Leu Phe Ile
510 515 520

cca tca tct aaa cta cta ttc ttg cgt tat cat aaa gaa aaa gag gtt 1938
Pro Ser Ser Lys Leu Leu Phe Leu Arg Tyr His Lys Glu Lys Glu Val
525 530 535

gtt gct gta gcc cat gct gtt tat caa gca gtg ctc agc ttg aag aat 1986
Val Ala Val Ala His Ala Val Tyr Gln Ala Val Leu Ser Leu Lys Asn
540 545 550

att cct gtt ttg gag act gcc tat aag tta ata ttg gga gaa atg act 2034
Ile Pro Val Leu Glu Thr Ala Tyr Lys Leu Ile Leu Gly Glu Met Thr
555 560 565

tgt gcc cta aac aac ctc cta cac agt cta caa ctt cct gag gcc tgt 2082
Cys Ala Leu Asn Asn Leu Leu His Ser Leu Gln Leu Pro Glu Ala Cys

570 575 580 585

tct gaa ata aaa cat gag gct ttt aag aat cat gtg ttc aat gta gac 2130
Ser Glu Ile Lys His Glu Ala Phe Lys Asn His Val Phe Asn Val Asp
590 595 600

aat gca aaa ttt gta gtt aaa ttt gac ctc agt gcc ctg act aca att 2178
Asn Ala Lys Phe Val Val Lys Phe Asp Leu Ser Ala Leu Thr Thr Ile
605 610 615

gga aat gcc aaa aac tca cta ata ggg atg tgg gcg cta tct cca act 2226
Gly Asn Ala Lys Asn Ser Leu Ile Gly Met Trp Ala Leu Ser Pro Thr
620 625 630

gtc ttt gca ctt ctg agt aag aat ctg atg att gtg cac agt gac ctg 2274
Val Phe Ala Leu Leu Ser Lys Asn Leu Met Ile Val His Ser Asp Leu
635 640 645

gct gtt cac ttc cct gcc att cag tat gct gtg ctc tac aca ttg tat 2322
Ala Val His Phe Pro Ala Ile Gln Tyr Ala Val Leu Tyr Thr Leu Tyr
650 655 660 665

tct cat tgt acc agg cat gat cac ttt atc tct agt agc ctc agt tct 2370
Ser His Cys Thr Arg His Asp His Phe Ile Ser Ser Ser Leu Ser Ser
670 675 680

gcc tct cct tct ttg ttt gat gga gct gtg att agc act gta act acg 2418
Ala Ser Pro Ser Leu Phe Asp Gly Ala Val Ile Ser Thr Val Thr Thr
685 690 695

gct aca aag aaa cat ttc tca att ata tta aat ctt ctg gga ata tta 2466
Ala Thr Lys Lys His Phe Ser Ile Ile Leu Asn Leu Leu Gly Ile Leu
700 705 710

ctt aag aaa gat aac ctt aac cag gac acg agg aaa ctg tta atg act 2514
Leu Lys Lys Asp Asn Leu Asn Gln Asp Thr Arg Lys Leu Leu Met Thr
715 720 725

tgg gct ttg gaa gca gct gtt tta atg agg aag tct gaa aca tac gca 2562
Trp Ala Leu Glu Ala Ala Val Leu Met Arg Lys Ser Glu Thr Tyr Ala
730 735 740 745

cct tta ttc tct ctt ccg tct ttc cat aaa ttt tgc aaa ggc ctt tta 2610
Pro Leu Phe Ser Leu Pro Ser Phe His Lys Phe Cys Lys Gly Leu Leu
750 755 760

gcc aac act ctc gtt gaa gat gtg aat atc tgt ctg cag gca tgc agc 2658
Ala Asn Thr Leu Val Glu Asp Val Asn Ile Cys Leu Gln Ala Cys Ser
765 770 775

agt cta cat gct ctg tcc tct tcc ttg cca gat gat ctt tta cag aga 2706
Ser Leu His Ala Leu Ser Ser Ser Leu Pro Asp Asp Leu Leu Gln Arg
780 785 790

tgt gtc gat gtt tgc cgt gtt caa cta gtg cac agt gga act cgt att 2754
Cys Val Asp Val Cys Arg Val Gln Leu Val His Ser Gly Thr Arg Ile
795 800 805

cga caa gca ttt gga aaa ctg ttg aaa tca att cct tta gat gtt gtc 2802
Arg Gln Ala Phe Gly Lys Leu Leu Lys Ser Ile Pro Leu Asp Val Val
810 815 820 825

cta agc aat aac aat cac aca gaa att caa gaa att tct tta gca tta 2850
Leu Ser Asn Asn Asn His Thr Glu Ile Gln Glu Ile Ser Leu Ala Leu
830 835 840

aga agt cac atg agt aaa gca cca agt aat aca ttc cac ccc caa gat 2898
Arg Ser His Met Ser Lys Ala Pro Ser Asn Thr Phe His Pro Gln Asp
845 850 855

ttc tct gat gtt att agt ttt att ttg tat ggg aac tct cat aga aca 2946
Phe Ser Asp Val Ile Ser Phe Ile Leu Tyr Gly Asn Ser His Arg Thr
860 865 870

ggg aag gac aat tgg ttg gaa aga ctg ttc tat agc tgc cag aga ctg 2994
Gly Lys Asp Asn Trp Leu Glu Arg Leu Phe Tyr Ser Cys Gln Arg Leu
875 880 885

gat aag cgt gac cag tca aca att cca cgc aat ctc ctg aag aca gat 3042
Asp Lys Arg Asp Gln Ser Thr Ile Pro Arg Asn Leu Leu Lys Thr Asp
890 895 900 905

gct gtc ctt tgg cag tgg gcc ata tgg gaa gct gca caa ttc act gtt 3090
Ala Val Leu Trp Gln Trp Ala Ile Trp Glu Ala Ala Gln Phe Thr Val
910 915 920

ctt tct aag ctg aga acc cca ctg ggc aga gct caa gac acc ttc cag 3138

Leu Ser Lys Leu Arg Thr Pro Leu Gly Arg Ala Gln Asp Thr Phe Gln
925 930 935

aca att gaa ggt atc att cga agt ctc gca gct cac aca tta aac cct 3186
Thr Ile Glu Gly Ile Ile Arg Ser Leu Ala Ala His Thr Leu Asn Pro
940 945 950

gat cag gat gtt agt cag tgg aca act gca gac aat gat gaa ggc cat 3234
Asp Gln Asp Val Ser Gln Trp Thr Thr Ala Asp Asn Asp Glu Gly His
955 960 965

ggt aac aac caa ctt aga ctt gtt ctt ctt ctg cag tat ctg gaa aat 3282
Gly Asn Asn Gln Leu Arg Leu Val Leu Leu Leu Gln Tyr Leu Glu Asn
970 975 980 985

ctg gag aaa tta atg tat aat gca tac gag gga tgt gct aat gca tta 3330
Leu Glu Lys Leu Met Tyr Asn Ala Tyr Glu Gly Cys Ala Asn Ala Leu
990 995 1000

act tca cct ccc aag gtc att aga act ttt ttc tat acc aat cgc caa 3378
Thr Ser Pro Pro Lys Val Ile Arg Thr Phe Phe Tyr Thr Asn Arg Gln
1005 1010 1015

act tgt cag gac tgg cta acg cgg att cga ctc tcc atc atg agg gta 3426
Thr Cys Gln Asp Trp Leu Thr Arg Ile Arg Leu Ser Ile Met Arg Val
1020 1025 1030

gga ttg ttg gca ggc cag cct gca gtg aca gtg aga cat ggc ttt gac 3474
Gly Leu Leu Ala Gly Gln Pro Ala Val Thr Val Arg His Gly Phe Asp

1035 1040 1045

ttg ctt aca gag atg aaa aca acc agc cta tct cag ggg aat gaa ttg 3522
Leu Leu Thr Glu Met Lys Thr Thr Ser Leu Ser Gln Gly Asn Glu Leu
1050 1055 1060 1065

gaa gta acc att atg atg gtg gta gaa gca tta tgt gaa ctt cat tgt 3570
Glu Val Thr Ile Met Met Val Val Glu Ala Leu Cys Glu Leu His Cys
1070 1075 1080

cct gaa gct ata cag gga att gct gtc tgg tca tca tct att gtt gga 3618
Pro Glu Ala Ile Gln Gly Ile Ala Val Trp Ser Ser Ser Ile Val Gly
1085 1090 1095

aaa aat ctt ctg tgg att aac tca gtg gct caa cag gct gaa ggg agg 3666
Lys Asn Leu Leu Trp Ile Asn Ser Val Ala Gln Gln Ala Glu Gly Arg
1100 1105 1110

ttt gaa aag gcc tct gtg gag tac cag gaa cac ctg tgt gcc atg aca 3714
Phe Glu Lys Ala Ser Val Glu Tyr Gln Glu His Leu Cys Ala Met Thr
1115 1120 1125

ggt gtt gat tgc tgc atc tcc agc ttt gac aaa tcg gtg ctc acc tta 3762
Gly Val Asp Cys Cys Ile Ser Ser Phe Asp Lys Ser Val Leu Thr Leu
1130 1135 1140 1145

gcc aat gct ggg cgt aac agt gcc agc ccg aaa cat tct ctg aat ggt 3810
Ala Asn Ala Gly Arg Asn Ser Ala Ser Pro Lys His Ser Leu Asn Gly
1150 1155 1160

gaa tcc aga aaa act gtg ctg tcc aaa ccg act gac tct tcc cct gag 3858
Glu Ser Arg Lys Thr Val Leu Ser Lys Pro Thr Asp Ser Ser Pro Glu
1165 1170 1175

gtt ata aat tat tta gga aat aaa gca tgt gag ttc tac atc tca att 3906
Val Ile Asn Tyr Leu Gly Asn Lys Ala Cys Glu Phe Tyr Ile Ser Ile
1180 1185 1190

gcc gat tgg gct gct gtg cag gaa tgg cag aac gct atc cat gac ttg 3954
Ala Asp Trp Ala Ala Val Gln Glu Trp Gln Asn Ala Ile His Asp Leu
1195 1200 1205

aaa aag agt acc agt agc act tcc ctc aac ctg aaa gct gac ttc aac 4002
Lys Lys Ser Thr Ser Ser Thr Ser Leu Asn Leu Lys Ala Asp Phe Asn
1210 1215 1220 1225

tat ata aaa tca tta agc agc ttt gag tct gga aaa ttt gtt gaa tgt 4050
Tyr Ile Lys Ser Leu Ser Ser Phe Glu Ser Gly Lys Phe Val Glu Cys
1230 1235 1240

acc gag cag tta gaa ttg tta cca gga gaa aat atc aat cta ctt gct 4098
Thr Glu Gln Leu Glu Leu Leu Pro Gly Glu Asn Ile Asn Leu Leu Ala
1245 1250 1255

gga gga tca aaa gaa aaa ata gac atg aaa aaa ctg ctt cct aac atg 4146
Gly Gly Ser Lys Glu Lys Ile Asp Met Lys Lys Leu Leu Pro Asn Met
1260 1265 1270

tta agt ccg gat ccg agg gaa ctt cag aaa tcc att gaa gtt caa ttg 4194
Leu Ser Pro Asp Pro Arg Glu Leu Gln Lys Ser Ile Glu Val Gln Leu
1275 1280 1285

tta aga agt tct gtt tgt ttg gca act gct tta aac ccg ata gaa caa 4242
Leu Arg Ser Ser Val Cys Leu Ala Thr Ala Leu Asn Pro Ile Glu Gln
1290 1295 1300 1305

gat cag aag tgg cag tct ata act gaa aat gtg gta aag tac ttg aag 4290
Asp Gln Lys Trp Gln Ser Ile Thr Glu Asn Val Val Lys Tyr Leu Lys
1310 1315 1320

caa aca tcc cgc atc gct att gga cct ctg aga ctt tct act tta aca 4338
Gln Thr Ser Arg Ile Ala Ile Gly Pro Leu Arg Leu Ser Thr Leu Thr
1325 1330 1335

gtt tca cag tct ttg cca gtt cta agt acc ttg cag ctg tat tgc tca 4386
Val Ser Gln Ser Leu Pro Val Leu Ser Thr Leu Gln Leu Tyr Cys Ser
1340 1345 1350

tct gct ttg gag aac aca gtt tct aac aga ctt tca aca gag gac tgt 4434
Ser Ala Leu Glu Asn Thr Val Ser Asn Arg Leu Ser Thr Glu Asp Cys
1355 1360 1365

ctt att cca ctc ttc agt gaa gct tta cgt tca tgt aaa cag cat gac 4482
Leu Ile Pro Leu Phe Ser Glu Ala Leu Arg Ser Cys Lys Gln His Asp
1370 1375 1380 1385

gtg agg cca tgg atg cag gca tta agg tat act atg tac cag aat cag 4530

Val Arg Pro Trp Met Gln Ala Leu Arg Tyr Thr Met Tyr Gln Asn Gln
1390 1395 1400

ttg ttg gag aaa att aaa gaa caa aca gtc cca att aga agc cat ctc 4578
Leu Leu Glu Lys Ile Lys Glu Gln Thr Val Pro Ile Arg Ser His Leu
1405 1410 1415

atg gaa tta ggt cta aca gca gca aaa ttt gct aga aaa cga ggg aat 4626
Met Glu Leu Gly Leu Thr Ala Ala Lys Phe Ala Arg Lys Arg Gly Asn
1420 1425 1430

gtg tcc ctt gca aca aga ctg ctg gca cag tgc agt gaa gtt cag ctg 4674
Val Ser Leu Ala Thr Arg Leu Leu Ala Gln Cys Ser Glu Val Gln Leu
1435 1440 1445

gga aag acc acc act gca cag gat tta gtc caa cat ttt aaa aaa cta 4722
Gly Lys Thr Thr Thr Ala Gln Asp Leu Val Gln His Phe Lys Lys Leu
1450 1455 1460 1465

tca acc caa ggt caa gtg gat gaa aaa tgg ggg ccc gaa ctt gat att 4770
Ser Thr Gln Gly Gln Val Asp Glu Lys Trp Gly Pro Glu Leu Asp Ile
1470 1475 1480

gaa aaa acc aaa ttg ctt tat aca gca ggc cag tca aca cat gca atg 4818
Glu Lys Thr Lys Leu Leu Tyr Thr Ala Gly Gln Ser Thr His Ala Met
1485 1490 1495

gaa atg ttg agt tct tgt gcc ata tct ttc tgc aag tct gtg aaa gct 4866
Glu Met Leu Ser Ser Cys Ala Ile Ser Phe Cys Lys Ser Val Lys Ala

1500 1505 1510

gaa tat gca gtt gct aaa tca att ctg aca ctg gct aaa tgg atc cag 4914
Glu Tyr Ala Val Ala Lys Ser Ile Leu Thr Leu Ala Lys Trp Ile Gln
1515 1520 1525

gca gaa tgg aaa gag att tca gga cag ctg aaa cag gtt tac aga gct 4962
Ala Glu Trp Lys Glu Ile Ser Gly Gln Leu Lys Gln Val Tyr Arg Ala
1530 1535 1540 1545

cag cac caa cag aac ttc aca ggt ctt tct act ttg tct aaa aac ata 5010
Gln His Gln Gln Asn Phe Thr Gly Leu Ser Thr Leu Ser Lys Asn Ile
1550 1555 1560

ctc act cta ata gaa ctg cca tct gtt aat acg atg gaa gaa gag tat 5058
Leu Thr Leu Ile Glu Leu Pro Ser Val Asn Thr Met Glu Glu Glu Tyr
1565 1570 1575

cct cgg atc gag agt gaa tct aca gtg cat att gga gtt gga gaa cct 5106
Pro Arg Ile Glu Ser Glu Ser Thr Val His Ile Gly Val Gly Glu Pro
1580 1585 1590

gac ttc att ttg gga cag ttg tat cac ctg tct tca gta cag gca cct 5154
Asp Phe Ile Leu Gly Gln Leu Tyr His Leu Ser Ser Val Gln Ala Pro
1595 1600 1605

gaa gta gcc aaa tct tgg gca gcg ttg gcc agc tgg gct tat agg tgg 5202
Glu Val Ala Lys Ser Trp Ala Ala Leu Ala Ser Trp Ala Tyr Arg Trp
1610 1615 1620 1625

ggc aga aag gtg gtt gac aat gcc agt cag gga gaa ggt gtt cgt ctg 5250
Gly Arg Lys Val Val Asp Asn Ala Ser Gln Gly Glu Gly Val Arg Leu
1630 1635 1640

ctg cct aga gaa aaa tct gaa gtt cag aat cta ctt cca gac act ata 5298
Leu Pro Arg Glu Lys Ser Glu Val Gln Asn Leu Leu Pro Asp Thr Ile
1645 1650 1655

act gag gaa gag aaa gag aga ata tat ggt att ctt gga cag gct gtg 5346
Thr Glu Glu Glu Lys Glu Arg Ile Tyr Gly Ile Leu Gly Gln Ala Val
1660 1665 1670

tgt cgg ccg gcg ggg att cag gat gaa gat ata aca ctt cag ata act 5394
Cys Arg Pro Ala Gly Ile Gln Asp Glu Asp Ile Thr Leu Gln Ile Thr
1675 1680 1685

gag agt gaa gac aac gaa gaa gat gac atg gtt gat gtt atc tgg cgt 5442
Glu Ser Glu Asp Asn Glu Glu Asp Asp Met Val Asp Val Ile Trp Arg
1690 1695 1700 1705

cag ttg ata tca agc tgc cca tgg ctt tca gaa ctt gat gaa agt gca 5490
Gln Leu Ile Ser Ser Cys Pro Trp Leu Ser Glu Leu Asp Glu Ser Ala
1710 1715 1720

act gaa gga gtt att aaa gtg tgg agg aaa gtt gta gat aga ata ttc 5538
Thr Glu Gly Val Ile Lys Val Trp Arg Lys Val Val Asp Arg Ile Phe
1725 1730 1735

agc ctg tac aaa ctc tct tgc agt gca tac ttt act ttc ctt aaa ctc 5586
Ser Leu Tyr Lys Leu Ser Cys Ser Ala Tyr Phe Thr Phe Leu Lys Leu
1740 1745 1750

aac gct ggt caa att cct tta gat gag gat gac cct agg ctg cat tta 5634
Asn Ala Gly Gln Ile Pro Leu Asp Glu Asp Asp Pro Arg Leu His Leu
1755 1760 1765

agt cac aga gtg gaa cag agc act gat gac atg att gtg atg gcc aca 5682
Ser His Arg Val Glu Gln Ser Thr Asp Asp Met Ile Val Met Ala Thr
1770 1775 1780 1785

ttg cgc ctg ctg cgg ttg ctc gtg aag cat gct ggt gag ctt cgg cag 5730
Leu Arg Leu Leu Arg Leu Leu Val Lys His Ala Gly Glu Leu Arg Gln
1790 1795 1800

tat ctg gag cac ggc ttg gag aca aca ccc act gca cca tgg agg gga 5778
Tyr Leu Glu His Gly Leu Glu Thr Thr Pro Thr Ala Pro Trp Arg Gly
1805 1810 1815

att att ccg caa ctt ttc tca cgc tta aac cac cct gaa gtg tat gtg 5826
Ile Ile Pro Gln Leu Phe Ser Arg Leu Asn His Pro Glu Val Tyr Val
1820 1825 1830

cgc caa agt att tgt aac ctt ctc tgc cgt gtg gct caa gat tcc cca 5874
Arg Gln Ser Ile Cys Asn Leu Leu Cys Arg Val Ala Gln Asp Ser Pro
1835 1840 1845

cat ctc ata ttg tat cct gca ata gtg ggt acc ata tcg ctt agt agt 5922

His Leu Ile Leu Tyr Pro Ala Ile Val Gly Thr Ile Ser Leu Ser Ser
1850 1855 1860 1865

gaa tcc cag gct tca gga aat aaa ttt tcc act gca att cca act tta 5970
Glu Ser Gln Ala Ser Gly Asn Lys Phe Ser Thr Ala Ile Pro Thr Leu
1870 1875 1880

ctt ggc aat att caa gga gaa gaa ttg ctg gtt tct gaa tgt gag gga 6018
Leu Gly Asn Ile Gln Gly Glu Glu Leu Leu Val Ser Glu Cys Glu Gly
1885 1890 1895

gga agt cct cct gca tct cag gat agc aat aag gat gaa cct aaa agt 6066
Gly Ser Pro Pro Ala Ser Gln Asp Ser Asn Lys Asp Glu Pro Lys Ser
1900 1905 1910

gga tta aat gaa gac caa gcc atg atg cag gat tgt tac agc aaa att 6114
Gly Leu Asn Glu Asp Gln Ala Met Met Gln Asp Cys Tyr Ser Lys Ile
1915 1920 1925

gta gat aag ctg tcc tct gca aac ccc acc atg gta tta cag gtt cag 6162
Val Asp Lys Leu Ser Ser Ala Asn Pro Thr Met Val Leu Gln Val Gln
1930 1935 1940 1945

atg ctc gtg gct gaa ctg cgc agg gtc act gtg ctc tgg gat gag ctc 6210
Met Leu Val Ala Glu Leu Arg Arg Val Thr Val Leu Trp Asp Glu Leu
1950 1955 1960

tgg ctg gga gtt ttg ctg caa caa cac atg tat gtc ctg aga cga att 6258
Trp Leu Gly Val Leu Leu Gln Gln His Met Tyr Val Leu Arg Arg Ile

1965 1970 1975

cag cag ctt gaa gat gag gtg aag aga gtc cag aac aac aac acc tta 6306
Gln Gln Leu Glu Asp Glu Val Lys Arg Val Gln Asn Asn Asn Thr Leu
1980 1985 1990

cgc aaa gaa gag aaa att gca atc atg agg gag agg cac aca gct ttg 6354
Arg Lys Glu Glu Lys Ile Ala Ile Met Arg Glu Arg His Thr Ala Leu
1995 2000 2005

atg aag ccc atc gta ttt gct ttg gag cat gtg agg agt atc aca gcg 6402
Met Lys Pro Ile Val Phe Ala Leu Glu His Val Arg Ser Ile Thr Ala
2010 2015 2020 2025

gct cct gca gaa aca cct cat gaa aaa tgg ttt cag gat aac tat ggt 6450
Ala Pro Ala Glu Thr Pro His Glu Lys Trp Phe Gln Asp Asn Tyr Gly
2030 2035 2040

gat gcc att gaa aat gcc cta gaa aaa ctg aag act cca ttg aac cct 6498
Asp Ala Ile Glu Asn Ala Leu Glu Lys Leu Lys Thr Pro Leu Asn Pro
2045 2050 2055

gca aag cct ggg agc agc tgg att cca ttt aaa gag ata atg cta agt 6546
Ala Lys Pro Gly Ser Ser Trp Ile Pro Phe Lys Glu Ile Met Leu Ser
2060 2065 2070

ttg caa cag aga gca cag aaa cgt gca agt tac atc ttg cgt ctt gaa 6594
Leu Gln Gln Arg Ala Gln Lys Arg Ala Ser Tyr Ile Leu Arg Leu Glu
2075 2080 2085

gaa atc agt cca tgg ttg gct gcc atg act aac act gaa att gct ctt 6642
Glu Ile Ser Pro Trp Leu Ala Ala Met Thr Asn Thr Glu Ile Ala Leu
2090 2095 2100 2105

cct ggg gaa gtc tca gcc aga gac act gtc aca atc cat agt gtg ggc 6690
Pro Gly Glu Val Ser Ala Arg Asp Thr Val Thr Ile His Ser Val Gly
2110 2115 2120

gga acc atc aca atc tta ccg act aaa acc aag cca aag aaa ctt ctc 6738
Gly Thr Ile Thr Ile Leu Pro Thr Lys Thr Lys Pro Lys Lys Leu Leu
2125 2130 2135

ttt ctt gga tca gat ggg aag agc tat cct tat ctt ttc aaa gga ctg 6786
Phe Leu Gly Ser Asp Gly Lys Ser Tyr Pro Tyr Leu Phe Lys Gly Leu
2140 2145 2150

gag gat tta cat ctg gat gag aga ata atg cag ttc cta tct att gtg 6834
Glu Asp Leu His Leu Asp Glu Arg Ile Met Gln Phe Leu Ser Ile Val
2155 2160 2165

aat acc atg ttt gct aca att aat cgc caa gaa aca ccc cgg ttc cat 6882
Asn Thr Met Phe Ala Thr Ile Asn Arg Gln Glu Thr Pro Arg Phe His
2170 2175 2180 2185

gct cga cac tat tct gta aca cca cta gga aca aga tca gga cta atc 6930
Ala Arg His Tyr Ser Val Thr Pro Leu Gly Thr Arg Ser Gly Leu Ile
2190 2195 2200

cag tgg gta gat gga gcc aca ccc tta ttt ggt ctt tac aaa cga tgg 6978
Gln Trp Val Asp Gly Ala Thr Pro Leu Phe Gly Leu Tyr Lys Arg Trp
2205 2210 2215

caa caa cgg gaa gct gcc tta caa gca caa aag gcc caa gat tcc tac 7026
Gln Gln Arg Glu Ala Ala Leu Gln Ala Gln Lys Ala Gln Asp Ser Tyr
2220 2225 2230

caa act cct cag aat cct gga att gta ccc cgt cct agt gaa ctt tat 7074
Gln Thr Pro Gln Asn Pro Gly Ile Val Pro Arg Pro Ser Glu Leu Tyr
2235 2240 2245

tac agt aaa att ggc cct gct ttg aaa aca gtt ggg ctt agc ctg gat 7122
Tyr Ser Lys Ile Gly Pro Ala Leu Lys Thr Val Gly Leu Ser Leu Asp
2250 2255 2260 2265

gtg tcc cgt cgg gat tgg cct ctt cat gta atg aag gca gta ttg gaa 7170
Val Ser Arg Arg Asp Trp Pro Leu His Val Met Lys Ala Val Leu Glu
2270 2275 2280

gag tta atg gag gcc aca ccc ccg aat ctc ctt gcc aaa gag ctc tgg 7218
Glu Leu Met Glu Ala Thr Pro Pro Asn Leu Leu Ala Lys Glu Leu Trp
2285 2290 2295

tca tct tgc aca aca cct gat gaa tgg tgg aga gtt acg cag tct tat 7266
Ser Ser Cys Thr Thr Pro Asp Glu Trp Trp Arg Val Thr Gln Ser Tyr
2300 2305 2310

gca aga tct act gca gtc atg tct atg gtt gga tac ata att ggc ctt 7314

Ala Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Ile Gly Leu
2315 2320 2325

gga gac aga cat ctg gat aat gtt ctt ata gat atg acg act gga gaa 7362
Gly Asp Arg His Leu Asp Asn Val Leu Ile Asp Met Thr Thr Gly Glu
2330 2335 2340 2345

gtt gtt cac ata gat tac aat gtt tgc ttt gaa aaa ggt aaa agc ctt 7410
Val Val His Ile Asp Tyr Asn Val Cys Phe Glu Lys Gly Lys Ser Leu
2350 2355 2360

aga gtt cct gag aaa gta cct ttt cga atg aca caa aac att gaa aca 7458
Arg Val Pro Glu Lys Val Pro Phe Arg Met Thr Gln Asn Ile Glu Thr
2365 2370 2375

gca ctg ggt gta act gga gta gaa ggt gta ttt agg ctt tca tgt gag 7506
Ala Leu Gly Val Thr Gly Val Glu Gly Val Phe Arg Leu Ser Cys Glu
2380 2385 2390

cag gtt tta cac att atg cgg cgt ggc aga gag acc ctg ctg acg ctg 7554
Gln Val Leu His Ile Met Arg Arg Gly Arg Glu Thr Leu Leu Thr Leu
2395 2400 2405

ctg gag gcc ttt gtg tac gac cct ctg gtg gac tgg aca gca gga ggc 7602
Leu Glu Ala Phe Val Tyr Asp Pro Leu Val Asp Trp Thr Ala Gly Gly
2410 2415 2420 2425

gag gct ggg ttt gct ggt gct gtc tat ggt gga ggt ggc cag cag gcc 7650
Glu Ala Gly Phe Ala Gly Ala Val Tyr Gly Gly Gly Gly Gln Gln Ala

2430 2435 2440

gag agc aag cag agc aag aga gag atg gag cga gag atc acc cgc agc 7698
Glu Ser Lys Gln Ser Lys Arg Glu Met Glu Arg Glu Ile Thr Arg Ser
2445 2450 2455

ctg ttt tct tct aga gta gct gag att aag gtg aac tgg ttt aag aat 7746
Leu Phe Ser Ser Arg Val Ala Glu Ile Lys Val Asn Trp Phe Lys Asn
2460 2465 2470

aga gat gag atg ctg gtt gtg ctt ccc aag ttg gac ggt agc tta gat 7794
Arg Asp Glu Met Leu Val Val Leu Pro Lys Leu Asp Gly Ser Leu Asp
2475 2480 2485

gaa tac cta agc ttg caa gag caa ctg aca gat gtg gaa aaa ctg cag 7842
Glu Tyr Leu Ser Leu Gln Glu Gln Leu Thr Asp Val Glu Lys Leu Gln
2490 2495 2500 2505

ggc aaa cta ctg gag gaa ata gag ttt cta gaa gga gct gaa ggg gtg 7890
Gly Lys Leu Leu Glu Glu Ile Glu Phe Leu Glu Gly Ala Glu Gly Val
2510 2515 2520

gat cat cct tct cat act ctg caa cac agg tat tct gag cac acc caa 7938
Asp His Pro Ser His Thr Leu Gln His Arg Tyr Ser Glu His Thr Gln
2525 2530 2535

cta cag act cag caa aga gct gtt cag gaa gca atc cag gtg aag ctg 7986
Leu Gln Thr Gln Gln Arg Ala Val Gln Glu Ala Ile Gln Val Lys Leu
2540 2545 2550

aat gaa ttt gaa caa tgg ata aca cat tat cag gct gca ttc aat aat 8034
Asn Glu Phe Glu Gln Trp Ile Thr His Tyr Gln Ala Ala Phe Asn Asn
2555 2560 2565

tta gaa gca aca cag ctt gca agc ttg ctt caa gag ata agc aca caa 8082
Leu Glu Ala Thr Gln Leu Ala Ser Leu Leu Gln Glu Ile Ser Thr Gln
2570 2575 2580 2585

atg gac ctt ggt cct cca agt tac gtg cca gca aca gcc ttt ctg cag 8130
Met Asp Leu Gly Pro Pro Ser Tyr Val Pro Ala Thr Ala Phe Leu Gln
2590 2595 2600

aat gct ggt cag gcc cac ttg att agc cag tgc gag cag ctg gag ggg 8178
Asn Ala Gly Gln Ala His Leu Ile Ser Gln Cys Glu Gln Leu Glu Gly
2605 2610 2615

gag gtt ggt gct ctc ctg cag cag agg cgc tcc gtg ctc cgt ggc tgt 8226
Glu Val Gly Ala Leu Leu Gln Gln Arg Arg Ser Val Leu Arg Gly Cys
2620 2625 2630

ctg gag caa ctg cat cac tat gca acc gtg gcc ctg cag tat ccg aag 8274
Leu Glu Gln Leu His His Tyr Ala Thr Val Ala Leu Gln Tyr Pro Lys
2635 2640 2645

gcc ata ttt cag aaa cat cga att gaa cag tgg aag acc tgg atg gaa 8322
Ala Ile Phe Gln Lys His Arg Ile Glu Gln Trp Lys Thr Trp Met Glu
2650 2655 2660 2665

gag ctc atc tgt aac acc aca gta gag cgt tgt caa gag ctc tat agg 8370
Glu Leu Ile Cys Asn Thr Thr Val Glu Arg Cys Gln Glu Leu Tyr Arg
2670 2675 2680

aaa tat gaa atg caa tat gct ccc cag cca ccc cca aca gtg tgt cag 8418
Lys Tyr Glu Met Gln Tyr Ala Pro Gln Pro Pro Pro Thr Val Cys Gln
2685 2690 2695

ttc atc act gcc act gaa atg acc ctg cag cga tac gca gca gac atc 8466
Phe Ile Thr Ala Thr Glu Met Thr Leu Gln Arg Tyr Ala Ala Asp Ile
2700 2705 2710

aac agc aga ctt att aga caa gtg gaa cgc ttg aaa cag gaa gct gtc 8514
Asn Ser Arg Leu Ile Arg Gln Val Glu Arg Leu Lys Gln Glu Ala Val
2715 2720 2725

act gtg cca gtt tgt gaa gat cag ttg aaa gaa att gaa cgt tgc att 8562
Thr Val Pro Val Cys Glu Asp Gln Leu Lys Glu Ile Glu Arg Cys Ile
2730 2735 2740 2745

aaa gtt ttc ctt cat gag aat gga gaa gaa gga tct ttg agt cta gca 8610
Lys Val Phe Leu His Glu Asn Gly Glu Glu Gly Ser Leu Ser Leu Ala
2750 2755 2760

agt gtt att att tct gcc ctt tgt acc ctt aca agg cgt aac ctg atg 8658
Ser Val Ile Ile Ser Ala Leu Cys Thr Leu Thr Arg Arg Asn Leu Met
2765 2770 2775

atg gaa ggt gca gcg tca agt gct gga gaa cag ctg gtt gat ctg act 8706

Met Glu Gly Ala Ala Ser Ser Ala Gly Glu Gln Leu Val Asp Leu Thr
2780 2785 2790

tct cgg gat gga gcc tgg ttc ttg gag gaa ctc tgc agt atg agc gga 8754
Ser Arg Asp Gly Ala Trp Phe Leu Glu Glu Leu Cys Ser Met Ser Gly
2795 2800 2805

aac gtc acc tgc ttg gtt cag tta ctg aag cag tgc cac ctg gtg cca 8802
Asn Val Thr Cys Leu Val Gln Leu Leu Lys Gln Cys His Leu Val Pro
2810 2815 2820 2825

cag gac tta gat atc ccg aac ccc atg gaa gcg tct gag aca gtt cac 8850
Gln Asp Leu Asp Ile Pro Asn Pro Met Glu Ala Ser Glu Thr Val His
2830 2835 2840

tta gcc aat gga gtg tat acc tca ctt cag gaa ttg aat tcg aat ttc 8898
Leu Ala Asn Gly Val Tyr Thr Ser Leu Gln Glu Leu Asn Ser Asn Phe
2845 2850 2855

cgg caa atc ata ttt cca gaa gca ctt cga tgt tta atg aaa ggg gaa 8946
Arg Gln Ile Ile Phe Pro Glu Ala Leu Arg Cys Leu Met Lys Gly Glu
2860 2865 2870

tac acg tta gaa agt atg ctg cat gaa ctg gac ggt ctt att gag cag 8994
Tyr Thr Leu Glu Ser Met Leu His Glu Leu Asp Gly Leu Ile Glu Gln
2875 2880 2885

acc acc gat ggc gtt ccc ctg cag act cta gtg gaa tct ctt cag gcc 9042
Thr Thr Asp Gly Val Pro Leu Gln Thr Leu Val Glu Ser Leu Gln Ala

2890 2895 2900 2905

tac tta aga aac gca gct atg gga ctg gaa gaa gaa aca cat gct cat 9090
Tyr Leu Arg Asn Ala Ala Met Gly Leu Glu Glu Glu Thr His Ala His
2910 2915 2920

tac atc gat gtt gcc aga cta cta cat gct cag tac ggt gaa tta atc 9138
Tyr Ile Asp Val Ala Arg Leu Leu His Ala Gln Tyr Gly Glu Leu Ile
2925 2930 2935

caa ccg aga aat ggt tca gtt gat gaa aca ccc aaa atg tca gct ggc 9186
Gln Pro Arg Asn Gly Ser Val Asp Glu Thr Pro Lys Met Ser Ala Gly
2940 2945 2950

cag atg ctt ttg gta gca ttc gat ggc atg ttt gct caa gtt gaa act 9234
Gln Met Leu Leu Val Ala Phe Asp Gly Met Phe Ala Gln Val Glu Thr
2955 2960 2965

gct ttc agc tta tta gtt gaa aag ttg aac aag atg gaa att ccc ata 9282
Ala Phe Ser Leu Leu Val Glu Lys Leu Asn Lys Met Glu Ile Pro Ile
2970 2975 2980 2985

gct tgg cga aag att gac atc ata agg gaa gcc agg agt act caa gtt 9330
Ala Trp Arg Lys Ile Asp Ile Ile Arg Glu Ala Arg Ser Thr Gln Val
2990 2995 3000

aat ttt ttt gat gat gat aat cac cgg cag gtg cta gaa gag att ttc 9378
Asn Phe Phe Asp Asp Asp Asn His Arg Gln Val Leu Glu Glu Ile Phe
3005 3010 3015

ttt cta aaa aga cta cag act att aag gag ttc ttc agg ctc tgt ggt 9426
Phe Leu Lys Arg Leu Gln Thr Ile Lys Glu Phe Phe Arg Leu Cys Gly
3020 3025 3030

acc ttt tct aaa aca ttg tca gga tca agt tca ctt gaa gat cag aat 9474
Thr Phe Ser Lys Thr Leu Ser Gly Ser Ser Ser Leu Glu Asp Gln Asn
3035 3040 3045

act gtg aat ggg cct gta cag att gtc aat gtg aaa acc ctt ttt aga 9522
Thr Val Asn Gly Pro Val Gln Ile Val Asn Val Lys Thr Leu Phe Arg
3050 3055 3060 3065

aac tct tgt ttc agt gaa gac caa atg gcc aaa cct atc aag gca ttc 9570
Asn Ser Cys Phe Ser Glu Asp Gln Met Ala Lys Pro Ile Lys Ala Phe
3070 3075 3080

aca gct gac ttt gtg agg cag ctc ttg ata ggg cta ccc aac caa gcc 9618
Thr Ala Asp Phe Val Arg Gln Leu Leu Ile Gly Leu Pro Asn Gln Ala
3085 3090 3095

ctc gga ctc aca ctg tgc agt ttt atc agt gct ctg ggt gta gac atc 9666
Leu Gly Leu Thr Leu Cys Ser Phe Ile Ser Ala Leu Gly Val Asp Ile
3100 3105 3110

att gct caa gta gag gca aag gac ttt ggt gcc gaa agc aaa gtt tct 9714
Ile Ala Gln Val Glu Ala Lys Asp Phe Gly Ala Glu Ser Lys Val Ser
3115 3120 3125

gtt gat gat ctc tgt aag aaa gcg gtg gaa cat aac atc cag ata ggg 9762
Val Asp Asp Leu Cys Lys Lys Ala Val Glu His Asn Ile Gln Ile Gly
3130 3135 3140 3145

aag ttc tct cag ctg gtt atg aac agg gca act gtg tta gca agt tct 9810
Lys Phe Ser Gln Leu Val Met Asn Arg Ala Thr Val Leu Ala Ser Ser
3150 3155 3160

tac gac act gcc tgg aag aag cat gac ttg gtg cga agg cta gaa acc 9858
Tyr Asp Thr Ala Trp Lys Lys His Asp Leu Val Arg Arg Leu Glu Thr
3165 3170 3175

agt att tct tct tgt aag aca agc ctg cag cgg gtt cag ctg cat att 9906
Ser Ile Ser Ser Cys Lys Thr Ser Leu Gln Arg Val Gln Leu His Ile
3180 3185 3190

gcc atg ttt cag tgg caa cat gaa gat cta ctt atc aat aga cca caa 9954
Ala Met Phe Gln Trp Gln His Glu Asp Leu Leu Ile Asn Arg Pro Gln
3195 3200 3205

gcc atg tca gtc aca cct ccc cca cgg tct gct atc cta acc agc atg 10002
Ala Met Ser Val Thr Pro Pro Pro Arg Ser Ala Ile Leu Thr Ser Met
3210 3215 3220 3225

aaa aag aag ctg cat acc ctg agc cag att gaa act tct att gcg aca 10050
Lys Lys Lys Leu His Thr Leu Ser Gln Ile Glu Thr Ser Ile Ala Thr
3230 3235 3240

gtt cag gag aag cta gct gca ctt gaa tca agt att gaa cag cga ctc 10098

Val Gln Glu Lys Leu Ala Ala Leu Glu Ser Ser Ile Glu Gln Arg Leu
3245 3250 3255

aag tgg gca ggt ggt gcc aac cct gca ttg gcc cct gta cta caa gat 10146
Lys Trp Ala Gly Gly Ala Asn Pro Ala Leu Ala Pro Val Leu Gln Asp
3260 3265 3270

ttt gaa gca acg ata gct gaa aga aga aat ctt gtc ctt aaa gag agc 10194
Phe Glu Ala Thr Ile Ala Glu Arg Arg Asn Leu Val Leu Lys Glu Ser
3275 3280 3285

caa aga gca agt cag gtc aca ttt ctc tgc agc aat atc att cat ttt 10242
Gln Arg Ala Ser Gln Val Thr Phe Leu Cys Ser Asn Ile Ile His Phe
3290 3295 3300 3305

gaa agt tta cga aca aga act gca gaa gcc tta aac ctg gat gcg gcg 10290
Glu Ser Leu Arg Thr Arg Thr Ala Glu Ala Leu Asn Leu Asp Ala Ala
3310 3315 3320

tta ttt gaa cta atc aag cga tgt cag cag atg tgt tcg ttt gca tca 10338
Leu Phe Glu Leu Ile Lys Arg Cys Gln Gln Met Cys Ser Phe Ala Ser
3325 3330 3335

cag ttt aac agt tca gtg tct gag tta gag ctt cgt tta tta cag aga 10386
Gln Phe Asn Ser Ser Val Ser Glu Leu Glu Leu Arg Leu Leu Gln Arg
3340 3345 3350

gtg gac act ggt ctt gaa cat cct att ggc agc tct gaa tgg ctt ttg 10434
Val Asp Thr Gly Leu Glu His Pro Ile Gly Ser Ser Glu Trp Leu Leu

3355 3360 3365

tca gca cac aaa cag ttg acc cag gat atg tct act cag agg gca att 10482
Ser Ala His Lys Gln Leu Thr Gln Asp Met Ser Thr Gln Arg Ala Ile
3370 3375 3380 3385

cag aca gag aaa gag cag cag ata gaa acg gtc tgt gaa aca att cag 10530
Gln Thr Glu Lys Glu Gln Gln Ile Glu Thr Val Cys Glu Thr Ile Gln
3390 3395 3400

aat ctg gtt gat aat ata aag act gtg ctc act ggt cat aac cga cag 10578
Asn Leu Val Asp Asn Ile Lys Thr Val Leu Thr Gly His Asn Arg Gln
3405 3410 3415

ctt gga gat gtc aaa cat ctc ttg aaa gct atg gct aag gat gaa gaa 10626
Leu Gly Asp Val Lys His Leu Leu Lys Ala Met Ala Lys Asp Glu Glu
3420 3425 3430

gct gct ctg gca gat ggt gaa gat gtt ccc tat gag aac agt gtt agg 10674
Ala Ala Leu Ala Asp Gly Glu Asp Val Pro Tyr Glu Asn Ser Val Arg
3435 3440 3445

cag ttt ttg ggt gaa tat aaa tca tgg caa gac aac att caa aca gtt 10722
Gln Phe Leu Gly Glu Tyr Lys Ser Trp Gln Asp Asn Ile Gln Thr Val
3450 3455 3460 3465

cta ttt aca tta gtc cag gct atg ggt cag gtt cga agt caa gaa cac 10770
Leu Phe Thr Leu Val Gln Ala Met Gly Gln Val Arg Ser Gln Glu His
3470 3475 3480

gtt gaa atg ctc cag gaa atc act ccc acc ttg aaa gaa ctg aaa aca 10818
Val Glu Met Leu Gln Glu Ile Thr Pro Thr Leu Lys Glu Leu Lys Thr
3485 3490 3495

caa agt cag agt atc tat aat aat tta gtg agt ttt gca tca ccc tta 10866
Gln Ser Gln Ser Ile Tyr Asn Asn Leu Val Ser Phe Ala Ser Pro Leu
3500 3505 3510

gtc acc gat gca aca aat gaa tgt tcg agt cca acg tca tct gct act 10914
Val Thr Asp Ala Thr Asn Glu Cys Ser Ser Pro Thr Ser Ser Ala Thr
3515 3520 3525

tat cag cca tcc ttc gct gca gca gtc cgg agt aac act ggc cag aag 10962
Tyr Gln Pro Ser Phe Ala Ala Ala Val Arg Ser Asn Thr Gly Gln Lys
3530 3535 3540 3545

act cag cct gat gtc atg tca cag aat gct aga aag ctg atc cag aaa 11010
Thr Gln Pro Asp Val Met Ser Gln Asn Ala Arg Lys Leu Ile Gln Lys
3550 3555 3560

aat ctt gct aca tca gct gat act cca cca agc acc gtt cca gga act 11058
Asn Leu Ala Thr Ser Ala Asp Thr Pro Pro Ser Thr Val Pro Gly Thr
3565 3570 3575

ggc aag agt gtt gct tgt agt cct aaa aag gca gtc aga gac cct aaa 11106
Gly Lys Ser Val Ala Cys Ser Pro Lys Lys Ala Val Arg Asp Pro Lys
3580 3585 3590

act ggg aaa gcg gtg caa gag aga aac tcc tat gca gtg agt gtg tgg 11154
Thr Gly Lys Ala Val Gln Glu Arg Asn Ser Tyr Ala Val Ser Val Trp
3595 3600 3605

aag aga gtg aaa gcc aag tta gag ggc cga gat gtt gat ccg aat agg 11202
Lys Arg Val Lys Ala Lys Leu Glu Gly Arg Asp Val Asp Pro Asn Arg
3610 3615 3620 3625

agg atg tca gtt gct gaa cag gtt gac tat gtc att aag gaa gca act 11250
Arg Met Ser Val Ala Glu Gln Val Asp Tyr Val Ile Lys Glu Ala Thr
3630 3635 3640

aat cta gat aac ttg gct cag ctg tat gaa ggt tgg aca gcc tgg gtg 11298
Asn Leu Asp Asn Leu Ala Gln Leu Tyr Glu Gly Trp Thr Ala Trp Val
3645 3650 3655

tga atggcaagac agtagatgag tctggttaag cgaggtcaga catccaccag 11351

aatcaactca gcctcaggca tccaaagcca caccacagtc ggtggtgatg caactggggg 11411

cttactctga ggaaacctag gaaatctcgg tgcactagga agtgaatccc gcaggacagc 11471

tgcactcagg gatacgccca acaccatggc ctgcaacccc agggtcaagg gtgaaggaaa 11531

gcaaagctca ccgcctgaac acggagattg tctttctgcc acagaacagc agcagacgtg 11591

tcgggaggtt agctgcggaa agaaatcggg atgccgcgga gcacagagtg atttggaact 11651

ccattccacc tgaccctgtg tgtacaatcc aggaaaaaaa caaaccccac tcagaaacag 11711

agaaaactgg ggtcgcgaag aaatcacagc caaggaagat ttgatgcatt cagattctcg 11771

tgtaacactt gttgcttggc aacagtactg gttgggttga ccagtaagta gaaaaaggct 11831

aaaggctatg cgatatgaat ttcagaaatg gactgaaaat ggagagctat gtaacagata 11891

cactacagta gaagaactta cttctgaaat gaagggaaaa aaaccacccc atcgttccct 11951

actcctcccc accacttacc cgttcccccc ttacctaatc tagtagatta gccatctttc 12011

aaattcactt ttatttcagt ccttatattt catatacttc cgtctcgatg ctgttaacaa 12071

cttctgataa catggaaaat tcaaggattg tttaaaggtc tgatgatcac acacaaaatg 12131

taattccggt tatttaagtc atttctgtga ttctatcatg tacagtttcc agaattgtca 12191

ctgtgcattc aaaagtaatg aatctaacag acatttgatt taatgtacac tccctttgc 12251

ttatagtgtg catttttttt ggaggtcatt caaattttcc ctcttctgtg atagctgtag 12311

tttctttcat agaaagtagc taatccagtg taatctttta ccttttttaaa aaccaagata 12371

gagtatctat tagagtttta cattgttgat gatagattaa caataaagtg atgttctggt 12431

ggaggtagac tgaaattttt ttaattcatg tttttcattt gatactttta atttacactt 12491

agtaaattaa aagttgttta atttacttgg cattttagga catgtacatg aaacagtgaa 12551

aatgagatcc accaacatct tttattaagt tcagttatta gtctgtgaag tgctttactt 12611

tttgcacaat tttaatagct tgctattcag taatacatta tagtgaattc atgatcaagg 12671

tttccttaaa tttagcattg catttcagta ctgactgtgt aagctaaatt gctgatccaa 12731

aataaaaacc cagactagaa tagggttctt aaaatcaagt atcaatacaa aatagaacac 12791

aattaaaatc ttaattgttg gctgggcaca gtggctcacg cctgtaatcc cagcactttg 12851

ggaggccgag gcgggcggat catgaggtta ggagagcgag accatcctgg ctaacacggt 12911

gaaaccccgt ctttactaaa atacaaaaaa aattagccgg gtgtggtggc gggcgcctgt 12971

agtcccagct actcgggagg ctgaggcagg agaatggcgt gaacccagga ggcggagctt 13031

gcagtgagcc gagattgtgc cactgcactc cagcctgggc aacagagcta gactctgtgt 13091

caaaaataaa tgactagat 13110

<210> 2

<211> 3657

<212> PRT

<213> Homo sapiens

<400> 2

Met Ser Arg Arg Ala Pro Gly Ser Arg Leu Ser Ser Gly Gly Thr Asn

1 5 10 15

Tyr Ser Arg Ser Trp Asn Asp Trp Gln Pro Arg Thr Asp Ser Ala Ser

20 25 30

Ala Asp Pro Gly Asn Leu Lys Tyr Ser Ser Ser Arg Asp Arg Gly Gly

35 40 45

Ser Ser Ser Tyr Gly Leu Gln Pro Ser Asn Ser Ala Val Val Ser Arg

50 55 60

Gln Arg His Asp Asp Thr Arg Val His Ala Asp Ile Gln Asn Asp Glu

65 70 75 80

Lys Gly Gly Tyr Ser Val Asn Gly Gly Ser Gly Glu Asn Thr Tyr Gly

85 90 95

Arg Lys Ser Leu Gly Gln Glu Leu Arg Val Asn Asn Val Thr Ser Pro

100 105 110

Glu Phe Thr Ser Val Gln His Gly Ser Arg Ala Leu Ala Thr Lys Asp

115 120 125

Met Arg Lys Ser Gln Glu Arg Ser Met Ser Tyr Ser Asp Glu Ser Arg

130 135 140

Leu Ser Asn Leu Leu Arg Arg Ile Thr Arg Glu Asp Asp Arg Asp Arg

145 150 155 160

Arg Leu Ala Thr Val Lys Gln Leu Lys Glu Phe Ile Gln Gln Pro Glu

165 170 175

Asn Lys Leu Val Leu Val Lys Gln Leu Asp Asn Ile Leu Ala Ala Val

180 185 190

His Asp Val Leu Asn Glu Ser Ser Lys Leu Leu Gln Glu Leu Arg Gln

195 200 205

Glu Gly Ala Cys Cys Leu Gly Leu Leu Cys Ala Ser Leu Ser Tyr Glu

210 215 220

Ala Glu Lys Ile Phe Lys Trp Ile Phe Ser Lys Phe Ser Ser Ser Ala

225 230 235 240

Lys Asp Glu Val Lys Leu Leu Tyr Leu Cys Ala Thr Tyr Lys Ala Leu
245 250 255
Glu Thr Val Gly Glu Lys Lys Ala Phe Ser Ser Val Met Gln Leu Val
260 265 270
Met Thr Ser Leu Gln Ser Ile Leu Glu Asn Val Asp Thr Pro Glu Leu
275 280 285
Leu Cys Lys Cys Val Lys Cys Ile Leu Leu Val Ala Arg Cys Tyr Pro
290 295 300
His Ile Phe Ser Thr Asn Phe Arg Asp Thr Val Asp Ile Leu Val Gly
305 310 315 320
Trp His Ile Asp His Thr Gln Lys Pro Ser Leu Thr Gln Gln Val Ser
325 330 335
Gly Trp Leu Gln Ser Leu Glu Pro Phe Trp Val Ala Asp Leu Ala Phe
340 345 350
Ser Thr Thr Leu Leu Gly Gln Phe Leu Glu Asp Met Glu Ala Tyr Ala
355 360 365
Glu Asp Leu Ser His Val Ala Ser Gly Glu Ser Val Asp Glu Asp Val
370 375 380
Pro Pro Pro Ser Val Ser Leu Pro Lys Leu Ala Ala Leu Leu Arg Val
385 390 395 400
Phe Ser Thr Val Val Arg Ser Ile Gly Glu Arg Phe Ser Pro Ile Arg
405 410 415
Gly Pro Pro Ile Thr Glu Ala Tyr Val Thr Asp Val Leu Tyr Arg Val
420 425 430
Met Arg Cys Val Thr Ala Ala Asn Gln Val Phe Phe Ser Glu Ala Val
435 440 445
Leu Thr Ala Ala Asn Glu Cys Val Gly Val Leu Leu Gly Ser Leu Asp
450 455 460
Pro Ser Met Thr Ile His Cys Asp Met Val Ile Thr Tyr Gly Leu Asp

465 470 475 480

Gln Leu Glu Asn Cys Gln Thr Cys Gly Thr Asp Tyr Ile Ile Ser Val

485 490 495

Leu Asn Leu Leu Thr Leu Ile Val Glu Gln Ile Asn Thr Lys Leu Pro

500 505 510

Ser Ser Phe Val Glu Lys Leu Phe Ile Pro Ser Ser Lys Leu Leu Phe

515 520 525

Leu Arg Tyr His Lys Glu Lys Glu Val Val Ala Val Ala His Ala Val

530 535 540

Tyr Gln Ala Val Leu Ser Leu Lys Asn Ile Pro Val Leu Glu Thr Ala

545 550 555 560

Tyr Lys Leu Ile Leu Gly Glu Met Thr Cys Ala Leu Asn Asn Leu Leu

565 570 575

His Ser Leu Gln Leu Pro Glu Ala Cys Ser Glu Ile Lys His Glu Ala

580 585 590

Phe Lys Asn His Val Phe Asn Val Asp Asn Ala Lys Phe Val Val Lys

595 600 605

Phe Asp Leu Ser Ala Leu Thr Thr Ile Gly Asn Ala Lys Asn Ser Leu

610 615 620

Ile Gly Met Trp Ala Leu Ser Pro Thr Val Phe Ala Leu Leu Ser Lys

625 630 635 640

Asn Leu Met Ile Val His Ser Asp Leu Ala Val His Phe Pro Ala Ile

645 650 655

Gln Tyr Ala Val Leu Tyr Thr Leu Tyr Ser His Cys Thr Arg His Asp

660 665 670

His Phe Ile Ser Ser Ser Leu Ser Ser Ala Ser Pro Ser Leu Phe Asp

675 680 685

Gly Ala Val Ile Ser Thr Val Thr Thr Ala Thr Lys Lys His Phe Ser

690 695 700

Ile Ile Leu Asn Leu Leu Gly Ile Leu Leu Lys Lys Asp Asn Leu Asn
705 710 715 720
Gln Asp Thr Arg Lys Leu Leu Met Thr Trp Ala Leu Glu Ala Ala Val
725 730 735
Leu Met Arg Lys Ser Glu Thr Tyr Ala Pro Leu Phe Ser Leu Pro Ser
740 745 750
Phe His Lys Phe Cys Lys Gly Leu Leu Ala Asn Thr Leu Val Glu Asp
755 760 765
Val Asn Ile Cys Leu Gln Ala Cys Ser Ser Leu His Ala Leu Ser Ser
770 775 780
Ser Leu Pro Asp Asp Leu Leu Gln Arg Cys Val Asp Val Cys Arg Val
785 790 795 800
Gln Leu Val His Ser Gly Thr Arg Ile Arg Gln Ala Phe Gly Lys Leu
805 810 815
Leu Lys Ser Ile Pro Leu Asp Val Val Leu Ser Asn Asn Asn His Thr
820 825 830
Glu Ile Gln Glu Ile Ser Leu Ala Leu Arg Ser His Met Ser Lys Ala
835 840 845
Pro Ser Asn Thr Phe His Pro Gln Asp Phe Ser Asp Val Ile Ser Phe
850 855 860
Ile Leu Tyr Gly Asn Ser His Arg Thr Gly Lys Asp Asn Trp Leu Glu
865 870 875 880
Arg Leu Phe Tyr Ser Cys Gln Arg Leu Asp Lys Arg Asp Gln Ser Thr
885 890 895
Ile Pro Arg Asn Leu Leu Lys Thr Asp Ala Val Leu Trp Gln Trp Ala
900 905 910
Ile Trp Glu Ala Ala Gln Phe Thr Val Leu Ser Lys Leu Arg Thr Pro
915 920 925
Leu Gly Arg Ala Gln Asp Thr Phe Gln Thr Ile Glu Gly Ile Ile Arg

930 935 940

Ser Leu Ala Ala His Thr Leu Asn Pro Asp Gln Asp Val Ser Gln Trp

945 950 955 960

Thr Thr Ala Asp Asn Asp Glu Gly His Gly Asn Asn Gln Leu Arg Leu

965 970 975

Val Leu Leu Leu Gln Tyr Leu Glu Asn Leu Glu Lys Leu Met Tyr Asn

980 985 990

Ala Tyr Glu Gly Cys Ala Asn Ala Leu Thr Ser Pro Pro Lys Val Ile

995 1000 1005

Arg Thr Phe Phe Tyr Thr Asn Arg Gln Thr Cys Gln Asp Trp Leu Thr

1010 1015 1020

Arg Ile Arg Leu Ser Ile Met Arg Val Gly Leu Leu Ala Gly Gln Pro

1025 1030 1035 1040

Ala Val Thr Val Arg His Gly Phe Asp Leu Leu Thr Glu Met Lys Thr

1045 1050 1055

Thr Ser Leu Ser Gln Gly Asn Glu Leu Glu Val Thr Ile Met Met Val

1060 1065 1070

Val Glu Ala Leu Cys Glu Leu His Cys Pro Glu Ala Ile Gln Gly Ile

1075 1080 1085

Ala Val Trp Ser Ser Ser Ile Val Gly Lys Asn Leu Leu Trp Ile Asn

1090 1095 1100

Ser Val Ala Gln Gln Ala Glu Gly Arg Phe Glu Lys Ala Ser Val Glu

1105 1110 1115 1120

Tyr Gln Glu His Leu Cys Ala Met Thr Gly Val Asp Cys Cys Ile Ser

1125 1130 1135

Ser Phe Asp Lys Ser Val Leu Thr Leu Ala Asn Ala Gly Arg Asn Ser

1140 1145 1150

Ala Ser Pro Lys His Ser Leu Asn Gly Glu Ser Arg Lys Thr Val Leu

1155 1160 1165

Ser Lys Pro Thr Asp Ser Ser Pro Glu Val Ile Asn Tyr Leu Gly Asn
1170 1175 1180

Lys Ala Cys Glu Phe Tyr Ile Ser Ile Ala Asp Trp Ala Ala Val Gln
1185 1190 1195 1200

Glu Trp Gln Asn Ala Ile His Asp Leu Lys Lys Ser Thr Ser Ser Thr
1205 1210 1215

Ser Leu Asn Leu Lys Ala Asp Phe Asn Tyr Ile Lys Ser Leu Ser Ser
1220 1225 1230

Phe Glu Ser Gly Lys Phe Val Glu Cys Thr Glu Gln Leu Glu Leu Leu
1235 1240 1245

Pro Gly Glu Asn Ile Asn Leu Leu Ala Gly Gly Ser Lys Glu Lys Ile
1250 1255 1260

Asp Met Lys Lys Leu Leu Pro Asn Met Leu Ser Pro Asp Pro Arg Glu
1265 1270 1275 1280

Leu Gln Lys Ser Ile Glu Val Gln Leu Leu Arg Ser Ser Val Cys Leu
1285 1290 1295

Ala Thr Ala Leu Asn Pro Ile Glu Gln Asp Gln Lys Trp Gln Ser Ile
1300 1305 1310

Thr Glu Asn Val Val Lys Tyr Leu Lys Gln Thr Ser Arg Ile Ala Ile
1315 1320 1325

Gly Pro Leu Arg Leu Ser Thr Leu Thr Val Ser Gln Ser Leu Pro Val
1330 1335 1340

Leu Ser Thr Leu Gln Leu Tyr Cys Ser Ser Ala Leu Glu Asn Thr Val
1345 1350 1355 1360

Ser Asn Arg Leu Ser Thr Glu Asp Cys Leu Ile Pro Leu Phe Ser Glu
1365 1370 1375

Ala Leu Arg Ser Cys Lys Gln His Asp Val Arg Pro Trp Met Gln Ala
1380 1385 1390

Leu Arg Tyr Thr Met Tyr Gln Asn Gln Leu Leu Glu Lys Ile Lys Glu

1395 1400 1405

Gln Thr Val Pro Ile Arg Ser His Leu Met Glu Leu Gly Leu Thr Ala

1410 1415 1420

Ala Lys Phe Ala Arg Lys Arg Gly Asn Val Ser Leu Ala Thr Arg Leu

1425 1430 1435 1440

Leu Ala Gln Cys Ser Glu Val Gln Leu Gly Lys Thr Thr Thr Ala Gln

1445 1450 1455

Asp Leu Val Gln His Phe Lys Lys Leu Ser Thr Gln Gly Gln Val Asp

1460 1465 1470

Glu Lys Trp Gly Pro Glu Leu Asp Ile Glu Lys Thr Lys Leu Leu Tyr

1475 1480 1485

Thr Ala Gly Gln Ser Thr His Ala Met Glu Met Leu Ser Ser Cys Ala

1490 1495 1500

Ile Ser Phe Cys Lys Ser Val Lys Ala Glu Tyr Ala Val Ala Lys Ser

1505 1510 1515 1520

Ile Leu Thr Leu Ala Lys Trp Ile Gln Ala Glu Trp Lys Glu Ile Ser

1525 1530 1535

Gly Gln Leu Lys Gln Val Tyr Arg Ala Gln His Gln Gln Asn Phe Thr

1540 1545 1550

Gly Leu Ser Thr Leu Ser Lys Asn Ile Leu Thr Leu Ile Glu Leu Pro

1555 1560 1565

Ser Val Asn Thr Met Glu Glu Glu Tyr Pro Arg Ile Glu Ser Glu Ser

1570 1575 1580

Thr Val His Ile Gly Val Gly Glu Pro Asp Phe Ile Leu Gly Gln Leu

1585 1590 1595 1600

Tyr His Leu Ser Ser Val Gln Ala Pro Glu Val Ala Lys Ser Trp Ala

1605 1610 1615

Ala Leu Ala Ser Trp Ala Tyr Arg Trp Gly Arg Lys Val Val Asp Asn

1620 1625 1630

Ala Ser Gln Gly Glu Gly Val Arg Leu Leu Pro Arg Glu Lys Ser Glu
1635 1640 1645

Val Gln Asn Leu Leu Pro Asp Thr Ile Thr Glu Glu Glu Lys Glu Arg
1650 1655 1660

Ile Tyr Gly Ile Leu Gly Gln Ala Val Cys Arg Pro Ala Gly Ile Gln
1665 1670 1675 1680

Asp Glu Asp Ile Thr Leu Gln Ile Thr Glu Ser Glu Asp Asn Glu Glu
1685 1690 1695

Asp Asp Met Val Asp Val Ile Trp Arg Gln Leu Ile Ser Ser Cys Pro
1700 1705 1710

Trp Leu Ser Glu Leu Asp Glu Ser Ala Thr Glu Gly Val Ile Lys Val
1715 1720 1725

Trp Arg Lys Val Val Asp Arg Ile Phe Ser Leu Tyr Lys Leu Ser Cys
1730 1735 1740

Ser Ala Tyr Phe Thr Phe Leu Lys Leu Asn Ala Gly Gln Ile Pro Leu
1745 1750 1755 1760

Asp Glu Asp Asp Pro Arg Leu His Leu Ser His Arg Val Glu Gln Ser
1765 1770 1775

Thr Asp Asp Met Ile Val Met Ala Thr Leu Arg Leu Leu Arg Leu Leu
1780 1785 1790

Val Lys His Ala Gly Glu Leu Arg Gln Tyr Leu Glu His Gly Leu Glu
1795 1800 1805

Thr Thr Pro Thr Ala Pro Trp Arg Gly Ile Ile Pro Gln Leu Phe Ser
1810 1815 1820

Arg Leu Asn His Pro Glu Val Tyr Val Arg Gln Ser Ile Cys Asn Leu
1825 1830 1835 1840

Leu Cys Arg Val Ala Gln Asp Ser Pro His Leu Ile Leu Tyr Pro Ala
1845 1850 1855

Ile Val Gly Thr Ile Ser Leu Ser Ser Glu Ser Gln Ala Ser Gly Asn

1860 1865 1870

Lys Phe Ser Thr Ala Ile Pro Thr Leu Leu Gly Asn Ile Gln Gly Glu

1875 1880 1885

Glu Leu Leu Val Ser Glu Cys Glu Gly Gly Ser Pro Pro Ala Ser Gln

1890 1895 1900

Asp Ser Asn Lys Asp Glu Pro Lys Ser Gly Leu Asn Glu Asp Gln Ala

1905 1910 1915 1920

Met Met Gln Asp Cys Tyr Ser Lys Ile Val Asp Lys Leu Ser Ser Ala

1925 1930 1935

Asn Pro Thr Met Val Leu Gln Val Gln Met Leu Val Ala Glu Leu Arg

1940 1945 1950

Arg Val Thr Val Leu Trp Asp Glu Leu Trp Leu Gly Val Leu Leu Gln

1955 1960 1965

Gln His Met Tyr Val Leu Arg Arg Ile Gln Gln Leu Glu Asp Glu Val

1970 1975 1980

Lys Arg Val Gln Asn Asn Asn Thr Leu Arg Lys Glu Glu Lys Ile Ala

1985 1990 1995 2000

Ile Met Arg Glu Arg His Thr Ala Leu Met Lys Pro Ile Val Phe Ala

2005 2010 2015

Leu Glu His Val Arg Ser Ile Thr Ala Ala Pro Ala Glu Thr Pro His

2020 2025 2030

Glu Lys Trp Phe Gln Asp Asn Tyr Gly Asp Ala Ile Glu Asn Ala Leu

2035 2040 2045

Glu Lys Leu Lys Thr Pro Leu Asn Pro Ala Lys Pro Gly Ser Ser Trp

2050 2055 2060

Ile Pro Phe Lys Glu Ile Met Leu Ser Leu Gln Gln Arg Ala Gln Lys

2065 2070 2075 2080

Arg Ala Ser Tyr Ile Leu Arg Leu Glu Glu Ile Ser Pro Trp Leu Ala

2085 2090 2095

Ala Met Thr Asn Thr Glu Ile Ala Leu Pro Gly Glu Val Ser Ala Arg
2100 2105 2110
Asp Thr Val Thr Ile His Ser Val Gly Gly Thr Ile Thr Ile Leu Pro
2115 2120 2125
Thr Lys Thr Lys Pro Lys Lys Leu Leu Phe Leu Gly Ser Asp Gly Lys
2130 2135 2140
Ser Tyr Pro Tyr Leu Phe Lys Gly Leu Glu Asp Leu His Leu Asp Glu
2145 2150 2155 2160
Arg Ile Met Gln Phe Leu Ser Ile Val Asn Thr Met Phe Ala Thr Ile
2165 2170 2175
Asn Arg Gln Glu Thr Pro Arg Phe His Ala Arg His Tyr Ser Val Thr
2180 2185 2190
Pro Leu Gly Thr Arg Ser Gly Leu Ile Gln Trp Val Asp Gly Ala Thr
2195 2200 2205
Pro Leu Phe Gly Leu Tyr Lys Arg Trp Gln Gln Arg Glu Ala Ala Leu
2210 2215 2220
Gln Ala Gln Lys Ala Gln Asp Ser Tyr Gln Thr Pro Gln Asn Pro Gly
2225 2230 2235 2240
Ile Val Pro Arg Pro Ser Glu Leu Tyr Tyr Ser Lys Ile Gly Pro Ala
2245 2250 2255
Leu Lys Thr Val Gly Leu Ser Leu Asp Val Ser Arg Arg Asp Trp Pro
2260 2265 2270
Leu His Val Met Lys Ala Val Leu Glu Glu Leu Met Glu Ala Thr Pro
2275 2280 2285
Pro Asn Leu Leu Ala Lys Glu Leu Trp Ser Ser Cys Thr Thr Pro Asp
2290 2295 2300
Glu Trp Trp Arg Val Thr Gln Ser Tyr Ala Arg Ser Thr Ala Val Met
2305 2310 2315 2320
Ser Met Val Gly Tyr Ile Ile Gly Leu Gly Asp Arg His Leu Asp Asn

2325 2330 2335

Val Leu Ile Asp Met Thr Thr Gly Glu Val Val His Ile Asp Tyr Asn

2340 2345 2350

Val Cys Phe Glu Lys Gly Lys Ser Leu Arg Val Pro Glu Lys Val Pro

2355 2360 2365

Phe Arg Met Thr Gln Asn Ile Glu Thr Ala Leu Gly Val Thr Gly Val

2370 2375 2380

Glu Gly Val Phe Arg Leu Ser Cys Glu Gln Val Leu His Ile Met Arg

2385 2390 2395 2400

Arg Gly Arg Glu Thr Leu Leu Thr Leu Leu Glu Ala Phe Val Tyr Asp

2405 2410 2415

Pro Leu Val Asp Trp Thr Ala Gly Gly Glu Ala Gly Phe Ala Gly Ala

2420 2425 2430

Val Tyr Gly Gly Gly Gly Gln Gln Ala Glu Ser Lys Gln Ser Lys Arg

2435 2440 2445

Glu Met Glu Arg Glu Ile Thr Arg Ser Leu Phe Ser Ser Arg Val Ala

2450 2455 2460

Glu Ile Lys Val Asn Trp Phe Lys Asn Arg Asp Glu Met Leu Val Val

2465 2470 2475 2480

Leu Pro Lys Leu Asp Gly Ser Leu Asp Glu Tyr Leu Ser Leu Gln Glu

2485 2490 2495

Gln Leu Thr Asp Val Glu Lys Leu Gln Gly Lys Leu Leu Glu Glu Ile

2500 2505 2510

Glu Phe Leu Glu Gly Ala Glu Gly Val Asp His Pro Ser His Thr Leu

2515 2520 2525

Gln His Arg Tyr Ser Glu His Thr Gln Leu Gln Thr Gln Gln Arg Ala

2530 2535 2540

Val Gln Glu Ala Ile Gln Val Lys Leu Asn Glu Phe Glu Gln Trp Ile

2545 2550 2555 2560

Thr His Tyr Gln Ala Ala Phe Asn Asn Leu Glu Ala Thr Gln Leu Ala
2565 2570 2575
Ser Leu Leu Gln Glu Ile Ser Thr Gln Met Asp Leu Gly Pro Pro Ser
2580 2585 2590
Tyr Val Pro Ala Thr Ala Phe Leu Gln Asn Ala Gly Gln Ala His Leu
2595 2600 2605
Ile Ser Gln Cys Glu Gln Leu Glu Gly Glu Val Gly Ala Leu Leu Gln
2610 2615 2620
Gln Arg Arg Ser Val Leu Arg Gly Cys Leu Glu Gln Leu His His Tyr
2625 2630 2635 2640
Ala Thr Val Ala Leu Gln Tyr Pro Lys Ala Ile Phe Gln Lys His Arg
2645 2650 2655
Ile Glu Gln Trp Lys Thr Trp Met Glu Glu Leu Ile Cys Asn Thr Thr
2660 2665 2670
Val Glu Arg Cys Gln Glu Leu Tyr Arg Lys Tyr Glu Met Gln Tyr Ala
2675 2680 2685
Pro Gln Pro Pro Pro Thr Val Cys Gln Phe Ile Thr Ala Thr Glu Met
2690 2695 2700
Thr Leu Gln Arg Tyr Ala Ala Asp Ile Asn Ser Arg Leu Ile Arg Gln
2705 2710 2715 2720
Val Glu Arg Leu Lys Gln Glu Ala Val Thr Val Pro Val Cys Glu Asp
2725 2730 2735
Gln Leu Lys Glu Ile Glu Arg Cys Ile Lys Val Phe Leu His Glu Asn
2740 2745 2750
Gly Glu Glu Gly Ser Leu Ser Leu Ala Ser Val Ile Ile Ser Ala Leu
2755 2760 2765
Cys Thr Leu Thr Arg Arg Asn Leu Met Met Glu Gly Ala Ala Ser Ser
2770 2775 2780
Ala Gly Glu Gln Leu Val Asp Leu Thr Ser Arg Asp Gly Ala Trp Phe

2785 2790 2795 2800

Leu Glu Glu Leu Cys Ser Met Ser Gly Asn Val Thr Cys Leu Val Gln

2805 2810 2815

Leu Leu Lys Gln Cys His Leu Val Pro Gln Asp Leu Asp Ile Pro Asn

2820 2825 2830

Pro Met Glu Ala Ser Glu Thr Val His Leu Ala Asn Gly Val Tyr Thr

2835 2840 2845

Ser Leu Gln Glu Leu Asn Ser Asn Phe Arg Gln Ile Ile Phe Pro Glu

2850 2855 2860

Ala Leu Arg Cys Leu Met Lys Gly Glu Tyr Thr Leu Glu Ser Met Leu

2865 2870 2875 2880

His Glu Leu Asp Gly Leu Ile Glu Gln Thr Thr Asp Gly Val Pro Leu

2885 2890 2895

Gln Thr Leu Val Glu Ser Leu Gln Ala Tyr Leu Arg Asn Ala Ala Met

2900 2905 2910

Gly Leu Glu Glu Glu Thr His Ala His Tyr Ile Asp Val Ala Arg Leu

2915 2920 2925

Leu His Ala Gln Tyr Gly Glu Leu Ile Gln Pro Arg Asn Gly Ser Val

2930 2935 2940

Asp Glu Thr Pro Lys Met Ser Ala Gly Gln Met Leu Leu Val Ala Phe

2945 2950 2955 2960

Asp Gly Met Phe Ala Gln Val Glu Thr Ala Phe Ser Leu Leu Val Glu

2965 2970 2975

Lys Leu Asn Lys Met Glu Ile Pro Ile Ala Trp Arg Lys Ile Asp Ile

2980 2985 2990

Ile Arg Glu Ala Arg Ser Thr Gln Val Asn Phe Phe Asp Asp Asp Asn

2995 3000 3005

His Arg Gln Val Leu Glu Glu Ile Phe Phe Leu Lys Arg Leu Gln Thr

3010 3015 3020

Ile Lys Glu Phe Phe Arg Leu Cys Gly Thr Phe Ser Lys Thr Leu Ser
3025 3030 3035 3040
Gly Ser Ser Ser Leu Glu Asp Gln Asn Thr Val Asn Gly Pro Val Gln
3045 3050 3055
Ile Val Asn Val Lys Thr Leu Phe Arg Asn Ser Cys Phe Ser Glu Asp
3060 3065 3070
Gln Met Ala Lys Pro Ile Lys Ala Phe Thr Ala Asp Phe Val Arg Gln
3075 3080 3085
Leu Leu Ile Gly Leu Pro Asn Gln Ala Leu Gly Leu Thr Leu Cys Ser
3090 3095 3100
Phe Ile Ser Ala Leu Gly Val Asp Ile Ile Ala Gln Val Glu Ala Lys
3105 3110 3115 3120
Asp Phe Gly Ala Glu Ser Lys Val Ser Val Asp Asp Leu Cys Lys Lys
3125 3130 3135
Ala Val Glu His Asn Ile Gln Ile Gly Lys Phe Ser Gln Leu Val Met
3140 3145 3150
Asn Arg Ala Thr Val Leu Ala Ser Ser Tyr Asp Thr Ala Trp Lys Lys
3155 3160 3165
His Asp Leu Val Arg Arg Leu Glu Thr Ser Ile Ser Ser Cys Lys Thr
3170 3175 3180
Ser Leu Gln Arg Val Gln Leu His Ile Ala Met Phe Gln Trp Gln His
3185 3190 3195 3200
Glu Asp Leu Leu Ile Asn Arg Pro Gln Ala Met Ser Val Thr Pro Pro
3205 3210 3215
Pro Arg Ser Ala Ile Leu Thr Ser Met Lys Lys Lys Leu His Thr Leu
3220 3225 3230
Ser Gln Ile Glu Thr Ser Ile Ala Thr Val Gln Glu Lys Leu Ala Ala
3235 3240 3245
Leu Glu Ser Ser Ile Glu Gln Arg Leu Lys Trp Ala Gly Gly Ala Asn

3250 3255 3260

Pro Ala Leu Ala Pro Val Leu Gln Asp Phe Glu Ala Thr Ile Ala Glu

3265 3270 3275 3280

Arg Arg Asn Leu Val Leu Lys Glu Ser Gln Arg Ala Ser Gln Val Thr

3285 3290 3295

Phe Leu Cys Ser Asn Ile Ile His Phe Glu Ser Leu Arg Thr Arg Thr

3300 3305 3310

Ala Glu Ala Leu Asn Leu Asp Ala Ala Leu Phe Glu Leu Ile Lys Arg

3315 3320 3325

Cys Gln Gln Met Cys Ser Phe Ala Ser Gln Phe Asn Ser Ser Val Ser

3330 3335 3340

Glu Leu Glu Leu Arg Leu Leu Gln Arg Val Asp Thr Gly Leu Glu His

3345 3350 3355 3360

Pro Ile Gly Ser Ser Glu Trp Leu Leu Ser Ala His Lys Gln Leu Thr

3365 3370 3375

Gln Asp Met Ser Thr Gln Arg Ala Ile Gln Thr Glu Lys Glu Gln Gln

3380 3385 3390

Ile Glu Thr Val Cys Glu Thr Ile Gln Asn Leu Val Asp Asn Ile Lys

3395 3400 3405

Thr Val Leu Thr Gly His Asn Arg Gln Leu Gly Asp Val Lys His Leu

3410 3415 3420

Leu Lys Ala Met Ala Lys Asp Glu Glu Ala Ala Leu Ala Asp Gly Glu

3425 3430 3435 3440

Asp Val Pro Tyr Glu Asn Ser Val Arg Gln Phe Leu Gly Glu Tyr Lys

3445 3450 3455

Ser Trp Gln Asp Asn Ile Gln Thr Val Leu Phe Thr Leu Val Gln Ala

3460 3465 3470

Met Gly Gln Val Arg Ser Gln Glu His Val Glu Met Leu Gln Glu Ile

3475 3480 3485

Thr Pro Thr Leu Lys Glu Leu Lys Thr Gln Ser Gln Ser Ile Tyr Asn
3490 3495 3500
Asn Leu Val Ser Phe Ala Ser Pro Leu Val Thr Asp Ala Thr Asn Glu
3505 3510 3515 3520
Cys Ser Ser Pro Thr Ser Ser Ala Thr Tyr Gln Pro Ser Phe Ala Ala
3525 3530 3535
Ala Val Arg Ser Asn Thr Gly Gln Lys Thr Gln Pro Asp Val Met Ser
3540 3545 3550
Gln Asn Ala Arg Lys Leu Ile Gln Lys Asn Leu Ala Thr Ser Ala Asp
3555 3560 3565
Thr Pro Pro Ser Thr Val Pro Gly Thr Gly Lys Ser Val Ala Cys Ser
3570 3575 3580
Pro Lys Lys Ala Val Arg Asp Pro Lys Thr Gly Lys Ala Val Gln Glu
3585 3590 3595 3600
Arg Asn Ser Tyr Ala Val Ser Val Trp Lys Arg Val Lys Ala Lys Leu
3605 3610 3615
Glu Gly Arg Asp Val Asp Pro Asn Arg Arg Met Ser Val Ala Glu Gln
3620 3625 3630
Val Asp Tyr Val Ile Lys Glu Ala Thr Asn Leu Asp Asn Leu Ala Gln
3635 3640 3645
Leu Tyr Glu Gly Trp Thr Ala Trp Val
3650 3655

<210> 3

<211> 22

<212> DNA

<213> Homo sapiens

<400> 3

agcgttatgt ttggtggaag aa 22

<210> 4

<211> 20

<212> DNA

<213> Homo sapiens

<400> 4

gcagctgtca acacagcctc 20

<210> 5

<211> 19

<212> DNA

<213> Homo sapiens

<400> 5

gatgtgtcga tgtttgccg 19

<210> 6

<211> 21

<212> DNA

<213> Homo sapiens

<400> 6

ttagcacatc cctcgtatgc a 21

<210> 7

<211> 15

<212> PRT

<213> Homo sapiens

<400> 7

Cys Asp Asn Leu Ala Gln Leu Tyr Glu Gly Trp Thr Ala Trp Val
1 5 10 15

<210> 8

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: A His tag sequence containing six histidine residues

<400> 8

Met Arg Gly Ser His His His His His His
1 5 10

[BRIEF DESCRIPTION OF THE DRAWINGS]

[Fig. 1] Figure 1 is a drawing showing the relationship between cDNA clones obtained in Example 1 and the novel base sequences and open reading frames obtained therefrom.

[Fig. 2] Figure 2 is a drawing showing the results of a

comparison between the human SMG-1 of the present invention and known proteins.

[Fig. 3] Figure 3 is a photograph, instead of a drawing, showing the results of autoradiography detection of the mRNA of human SMG-1 in various human cell lines.

[Fig. 4] Figure 4 is a drawing showing antigen sites used for preparing antibodies against human SMG-1.

[Fig. 5] Figure 5 is a photograph, instead of a drawing, showing the results of Western blotting for the HeLa cell lysate.

[Fig. 6] Figure 6 is a photograph, instead of a drawing, showing the results of Western blotting for various animal cell lysates.

[Fig. 7] Figure 7 is a photograph, instead of a drawing, showing the results of Western blotting for cell lysates derived from various animal tissues.

[Fig. 8] Figure 8 is a photograph, instead of a drawing, showing results of Western blotting and the results of confirmation of protein kinase activity, with respect to the immunoprecipitate derived from the HeLa cell lysate.

[Fig. 9] Figure 9 is a photograph, instead of a drawing, showing the expression of 6H-hSMG-1 and 6H-hSMG-1 (DA) and results of confirmation of in vitro protein kinase activity.

[Fig. 10] Figure 10 is a drawing schematically showing the structure of a reporter gene plasmid.

[Fig. 11] Figure 11 is a photograph, instead of a drawing, showing the results of evaluation of the amount of accumulation of reporter mRNA by Northern blotting.

[Fig. 12] Figure 12 is a photograph, instead of a drawing, showing representative examples of the results of confirmation of the effects of 6H-hSMG-1 and 6H-hSMG-1 (DA) on the accumulation of reporter mRNA.

[Fig. 13] Figure 13 is a graph of the results of statistical processing of the results of confirmation of the effects of 6H-hSMG-1 and 6H-hSMG-1 (DA) on the accumulation of reporter mRNA.

[Fig. 14] Figure 14 is a photograph, instead of a drawing, showing representative examples of the results of confirmation of the effects of 6H-hSMG-1 and 6H-hSMG-1 (DA)

on the accumulation of reporter mRNA in the presence of doxycycline where BGG-WT was used as a reporter mRNA.

[Fig. 15] Figure 15 is a graph of the results of a graphing of the results shown in Figure 14.

[Fig. 16] Figure 16 is a photograph, instead of a drawing, showing the results of confirmation of the effects of 6H-hSMG-1 and 6H-hSMG-1 (DA) on the accumulation of mRNA in the presence of doxycycline where BGG-39PTC was used as the reporter mRNA.

[Fig. 17] Figure 17 is a graph of the results of a graphing of the results shown in Figure 14.

[Fig. 18] Figure 18 is a photograph, instead of a drawing, showing the results of confirmation of the phosphorylation of full-length hUpf1/SMG-2 fusion protein by 6H-hSMG-1.

[Fig. 19] Figure 19 is a drawing schematically showing the structure of hUpf1/SMG-2 partial fragments used in Example 9(2).

[Fig. 20] Figure 20 is a photograph, instead of a drawing, showing the results of confirmation of the phosphorylation in fusion proteins of hUpf1/SMG-2 partial fragments by 6H-hSMG-1.

[Fig. 21] Figure 21 is a drawing schematically showing the structure of hUpf1/SMG-2 partial peptides used in Example 9(3).

[Fig. 22] Figure 22 is a photograph, instead of a drawing, showing the results of confirmation of the phosphorylation in fusion proteins of hUpf1/SMG-2 partial peptides by 6H-hSMG-1.

[Fig. 23] Figure 23 is a photograph, instead of a drawing, showing the results of confirmation of the phosphorylation of hUpf1/SMG-2 in the presence of okadaic acid in vivo.

[Fig. 24] Figure 24 is a photograph, instead of a drawing, showing the results of confirmation of the phosphorylation of hUpf1/SMG-2 in vivo using alkaline phosphatase.

[Fig. 25] Figure 25 is a photograph, instead of a drawing, showing the results of confirmation of the

phosphorylation of HA-hUpf1/SMG-2 in the case of an overexpression of 6H-hSMG-1 or 6H-hSMG-1 (DA).

[Fig. 26] Figure 26 is a graph showing the inhibitory effect of wortmannin on the kinase activity of 6H-hSMG-1.

[Fig. 27] Figure 27 is a graph showing the inhibitory effect of caffeine on the kinase activity of 6H-hSMG-1.

[Fig. 28] Figure 28 is a photograph, instead of a drawing, showing the results of confirmation of the inhibition by SMG-1 inhibitors on the phosphorylation of hUpf1/SMG-2 in the cell.

[Fig. 29] Figure 29 is a photograph, instead of a drawing, showing the stabilization of the endogenous PTC containing BGG gene product by SMG-1 inhibitors.

[Fig. 30] Figure 30 is a drawing schematically showing the structure of the p53 gene and the PTC mutations in the cell lines calu6 and N417.

[Fig. 31] Figure 31 is a photograph, instead of a drawing, showing the stabilization of the endogenous PTCp53 gene product by the SMG-1 inhibitor (wortmannin).

[Fig. 32] Figure 32 is a photograph, instead of a drawing, showing the stabilization of the endogenous PTCp53 gene product by various concentrations of SMG-1 inhibitors (wortmannin or caffeine).

[DOCUMENT NAME] Abstract

[ABSTRACT]

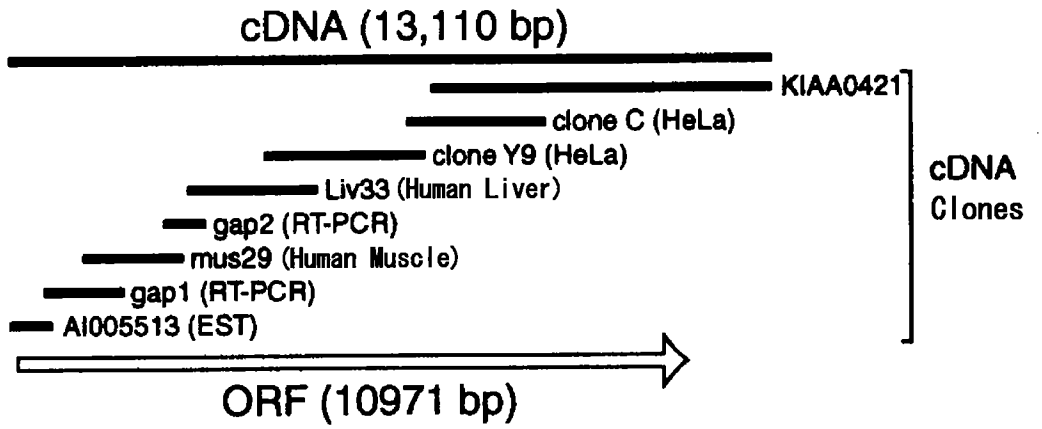
[OBJECT] A novel polypeptide, which is useful in constructing a screening system for agents of treating a disease caused by a premature translation termination codon generated by a nonsense mutation, and a novel polynucleotide encoding the polypeptide are provided.

[MEANS FOR SOLUTION] The polypeptide is SMG-1, a protein included in the phosphatidyl inositol kinase related kinase family.

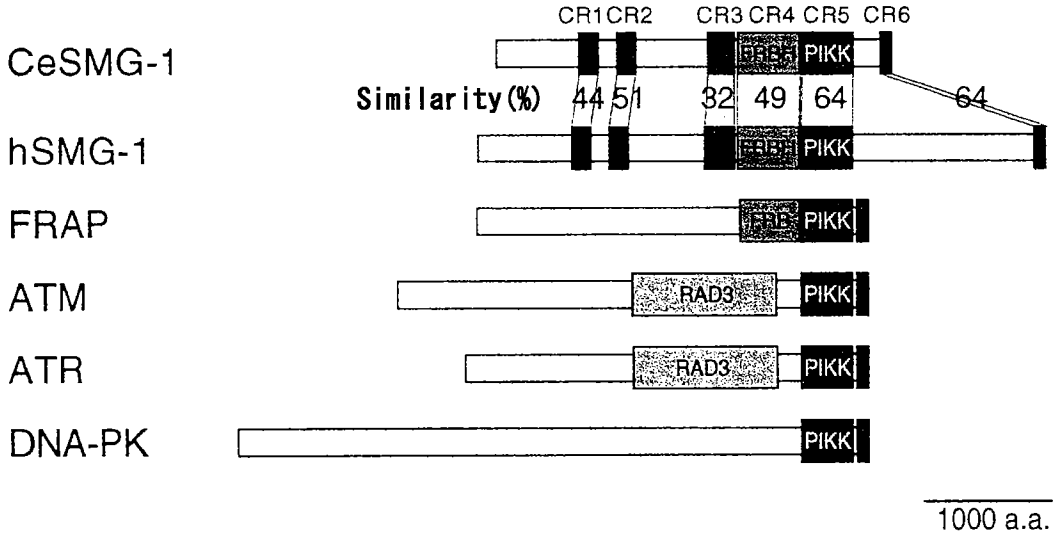
[SELECTED DRAWINGS] None

[DOCUMENT NAME] Drawings

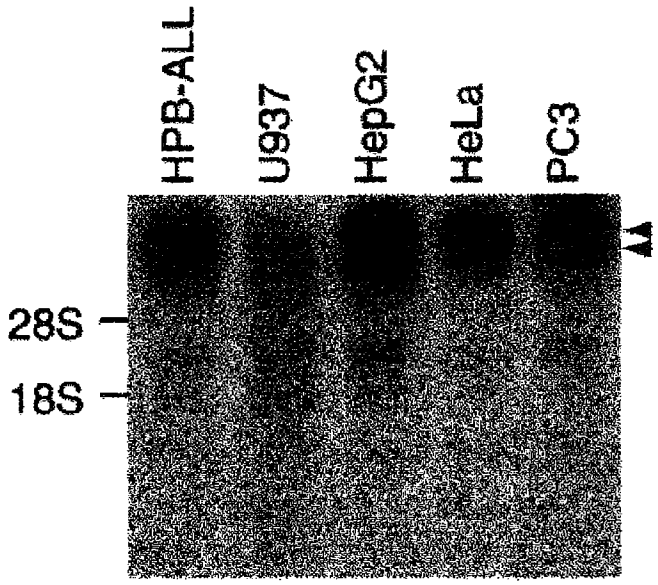
[Figure 1]



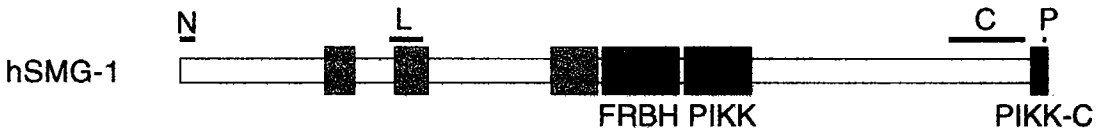
[Figure 2]



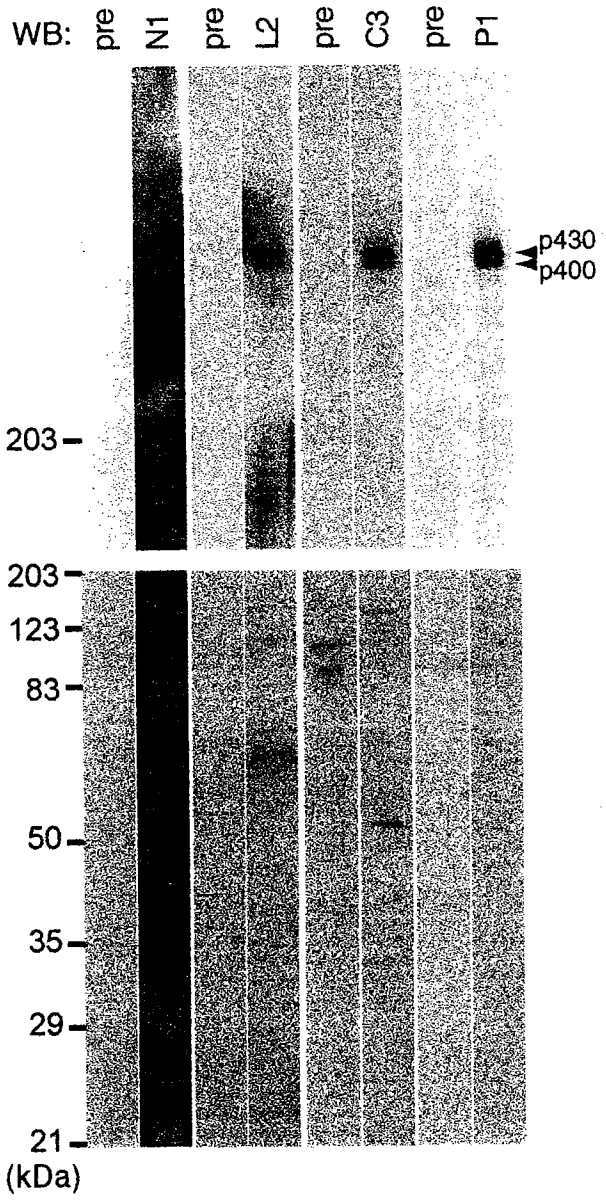
[Figure 3]



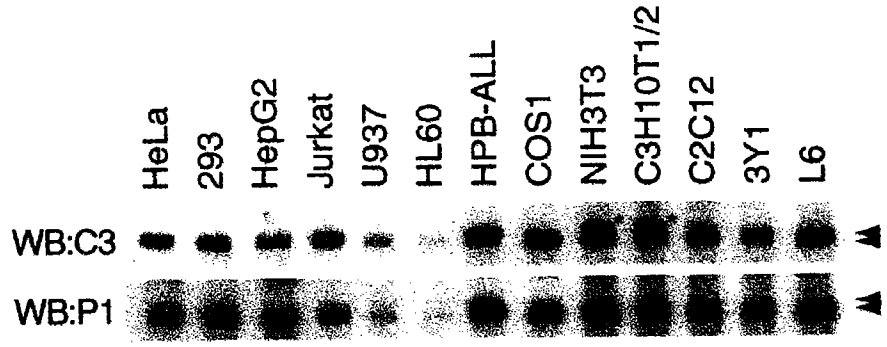
[Figure 4]



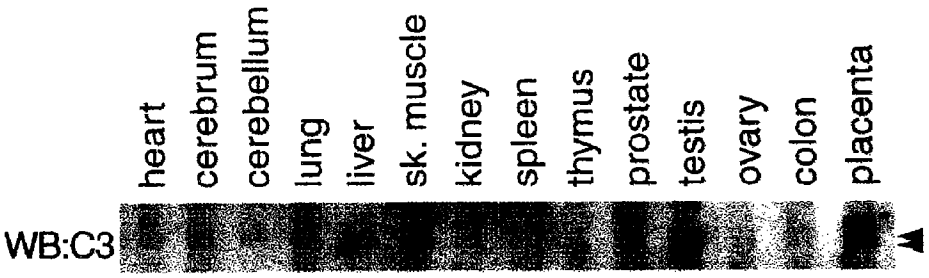
[Figure 5]



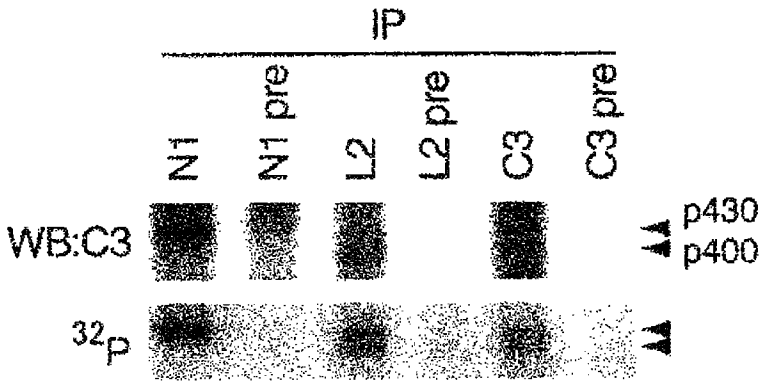
[Figure 6]



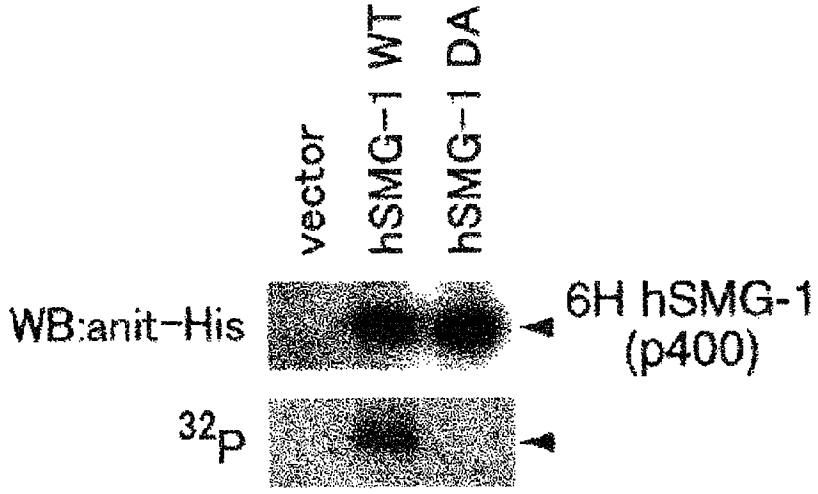
[Figure 7]



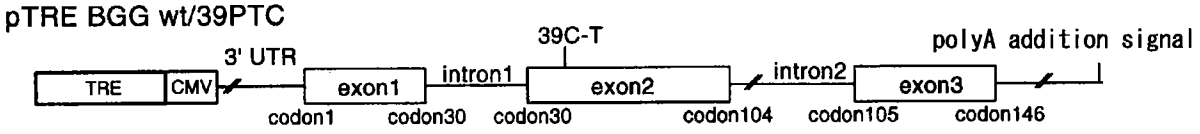
[Figure 8]



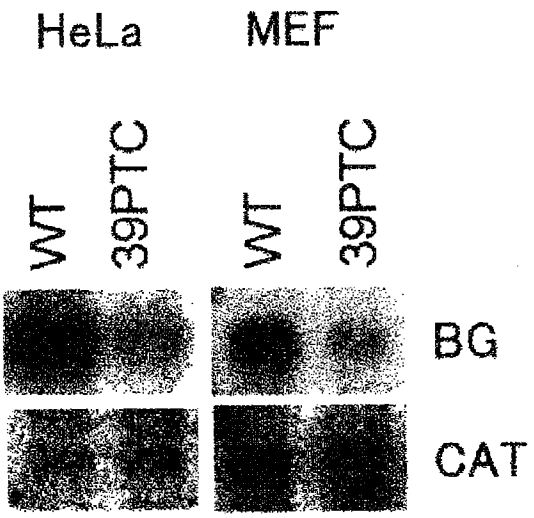
[Figure 9]



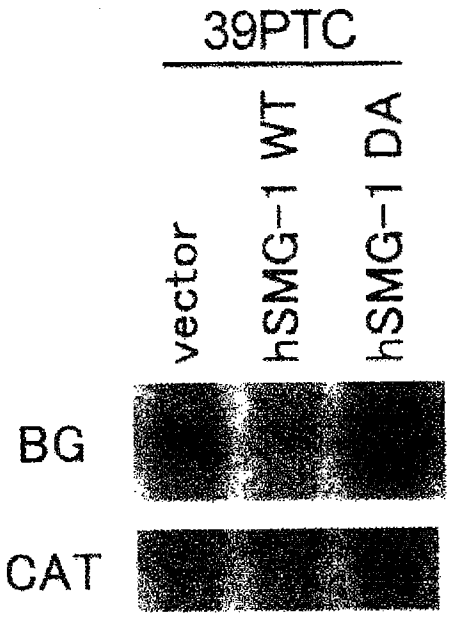
[Figure 10]



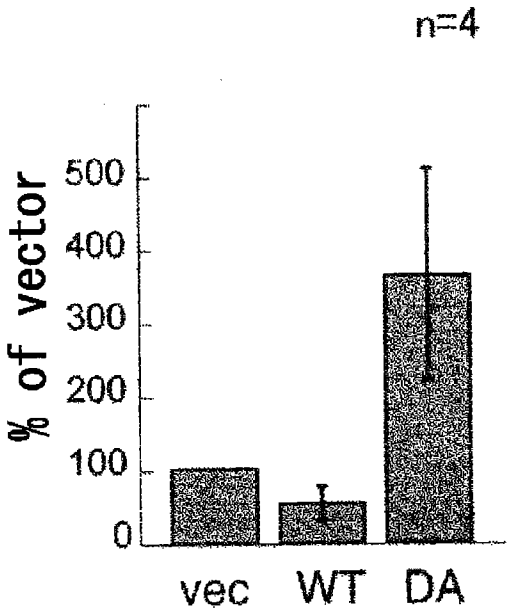
[Figure 11]



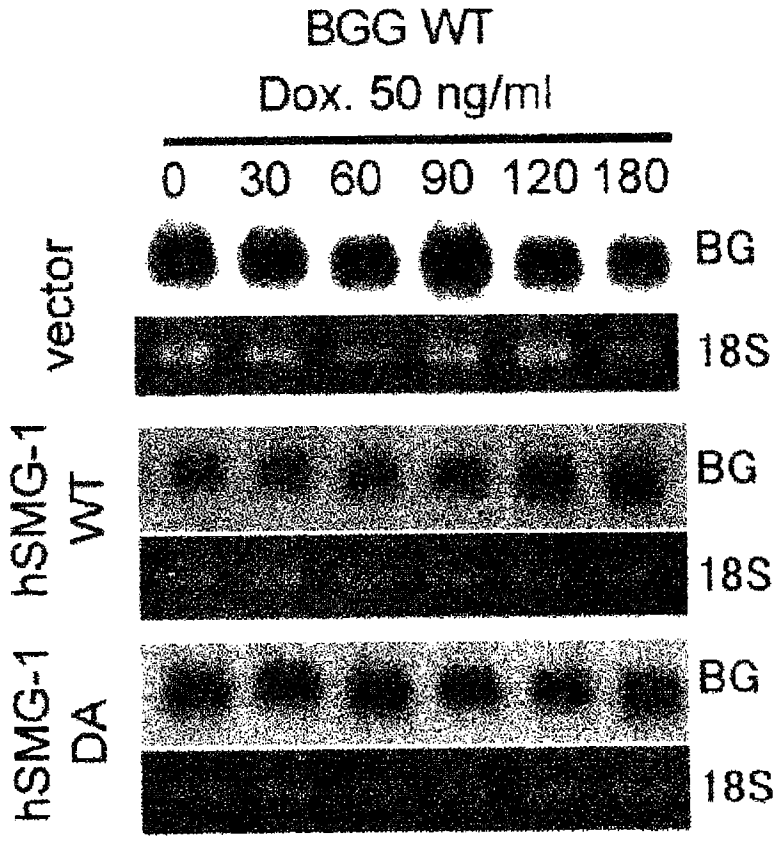
[Figure 12]



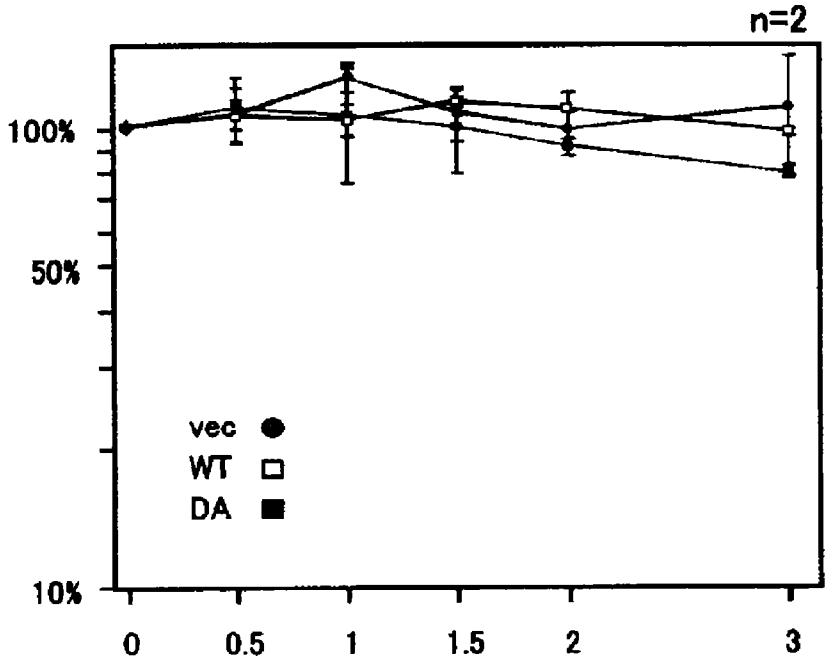
[Figure 13]



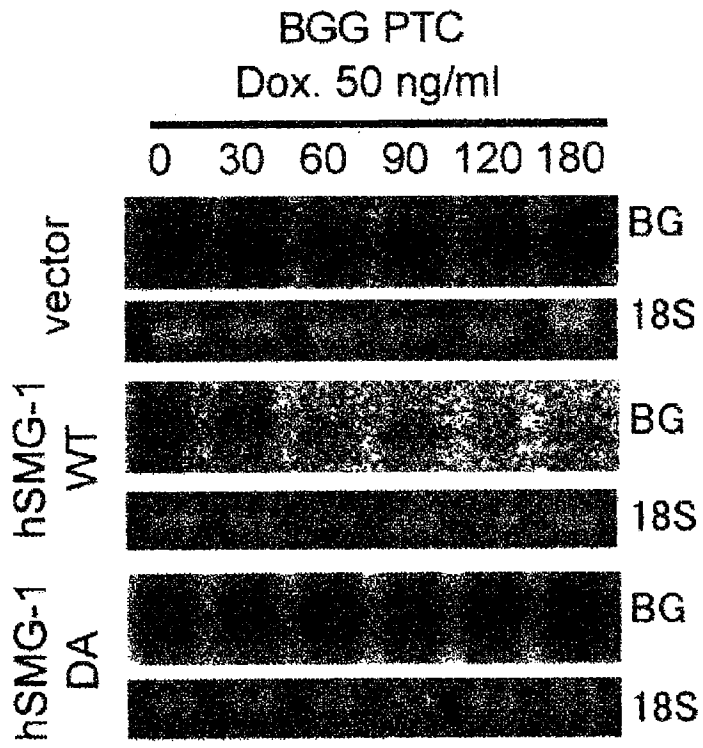
[Figure 14]



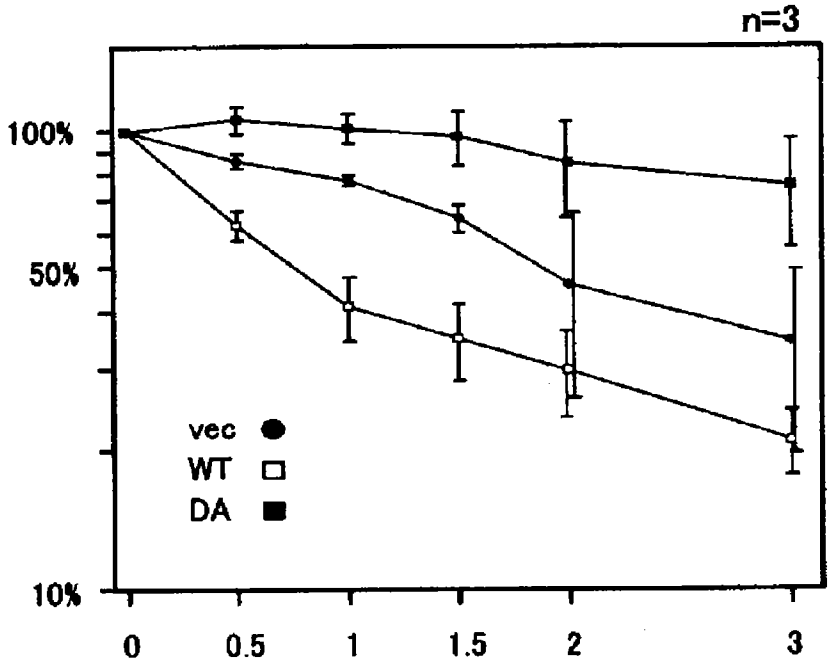
[Figure 15]



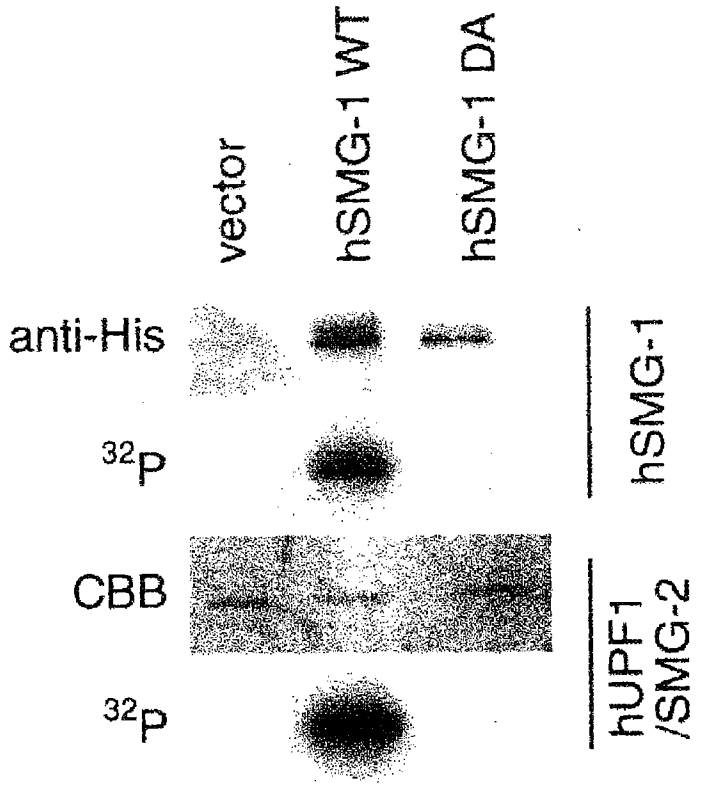
[Figure 16]



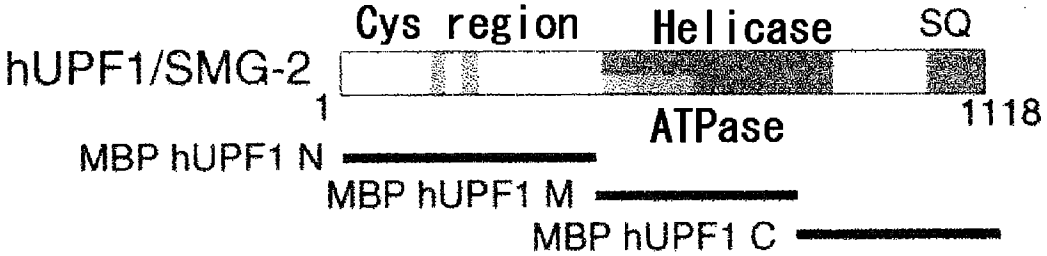
[Figure 17]



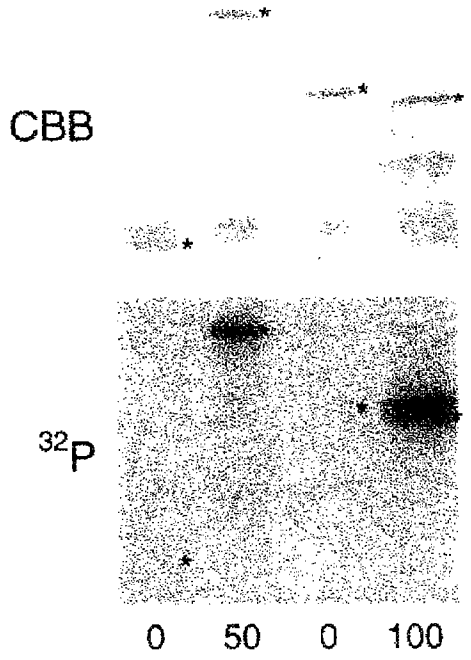
[Figure 18]



[Figure 19]



[Figure 20]

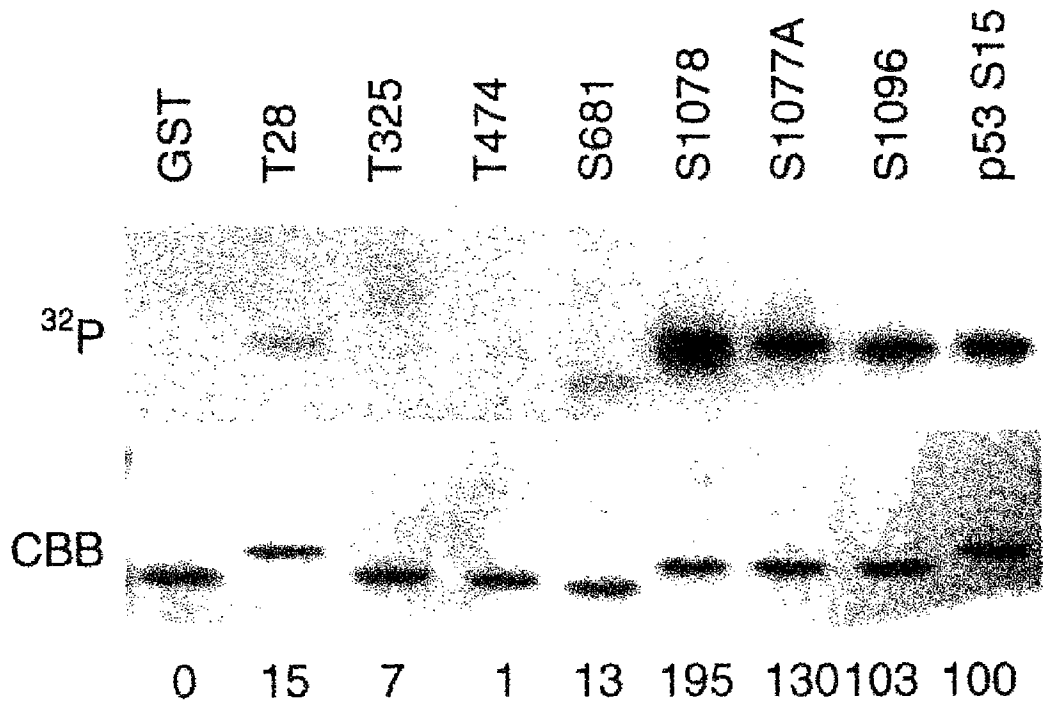


[Figure 21]

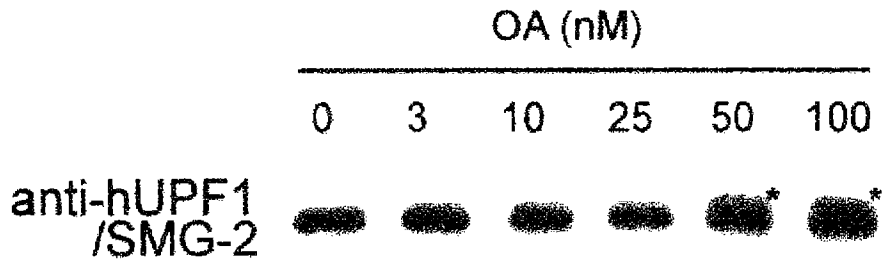
hUPF1/SMG-2 peptides

T28	E	L	L	G	A	D	T	Q	G	S	E	F	E	F
T325	K	L	K	E	S	Q	T	Q	D	N	I	T	V	R
S474	L	P	D	L	N	H	S	Q	V	Y	A	V	K	T
S681	A	A	K	A	G	L	S	Q	S	L	F	E	R	L
S1078	L	S	Q	P	E	L	S	Q	D	S	Y	L	G	D
S1096	Q	I	D	V	A	L	S	Q	D	S	T	Y	Q	G
p53 S15	S	V	E	P	P	L	S	Q	E	T	F	S	D	L

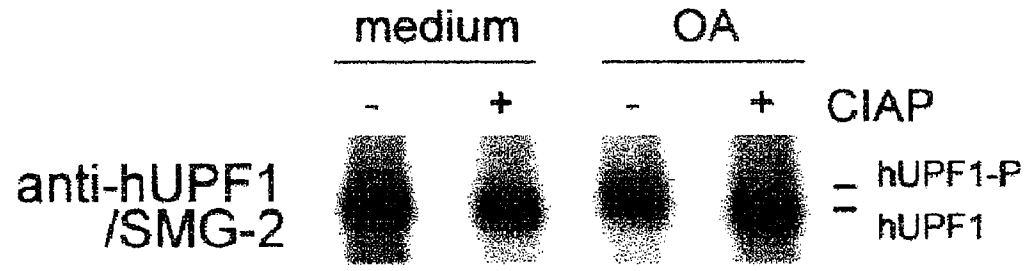
[Figure 22]



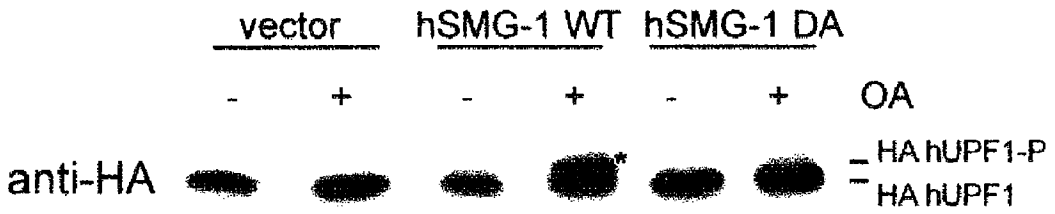
[Figure 23]



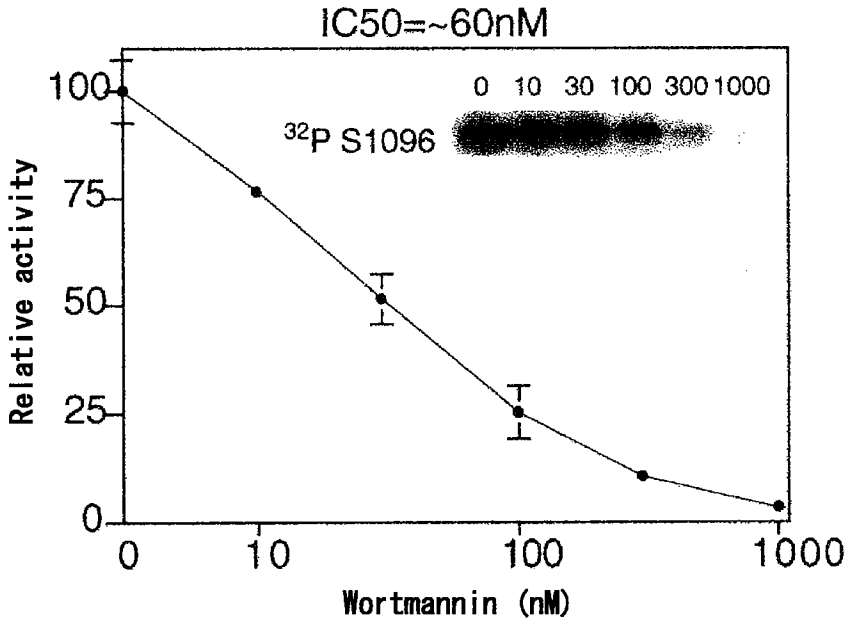
[Figure 24]



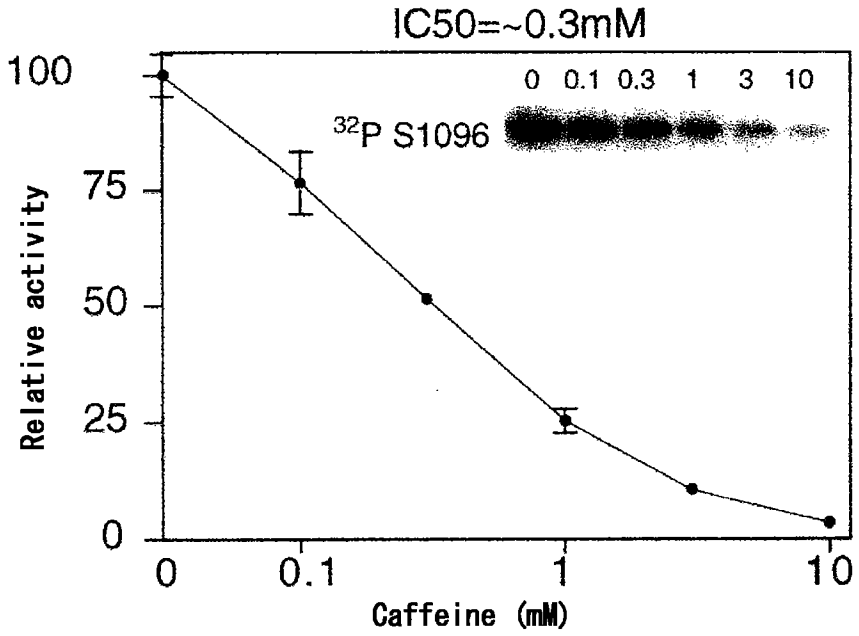
[Figure 25]



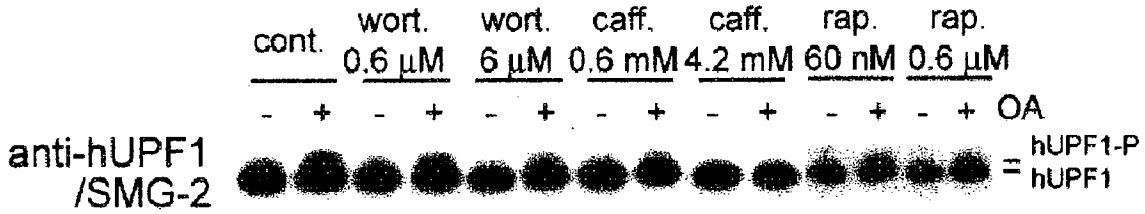
[Figure 26]



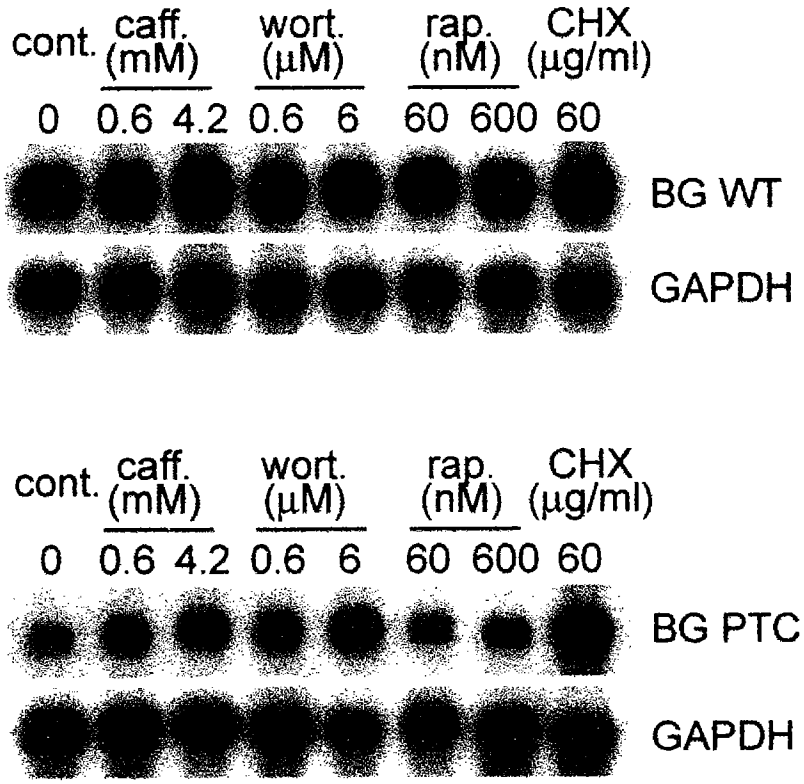
[Figure 27]



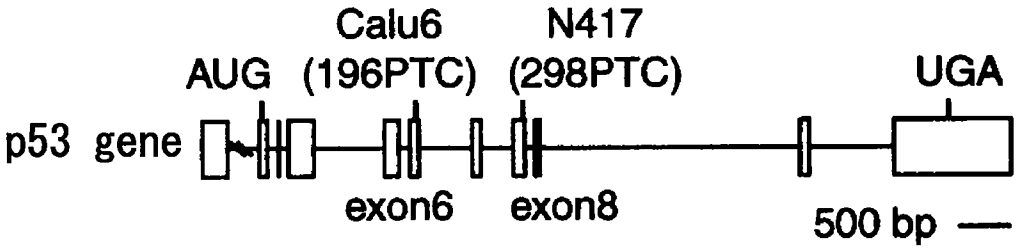
[Figure 28]



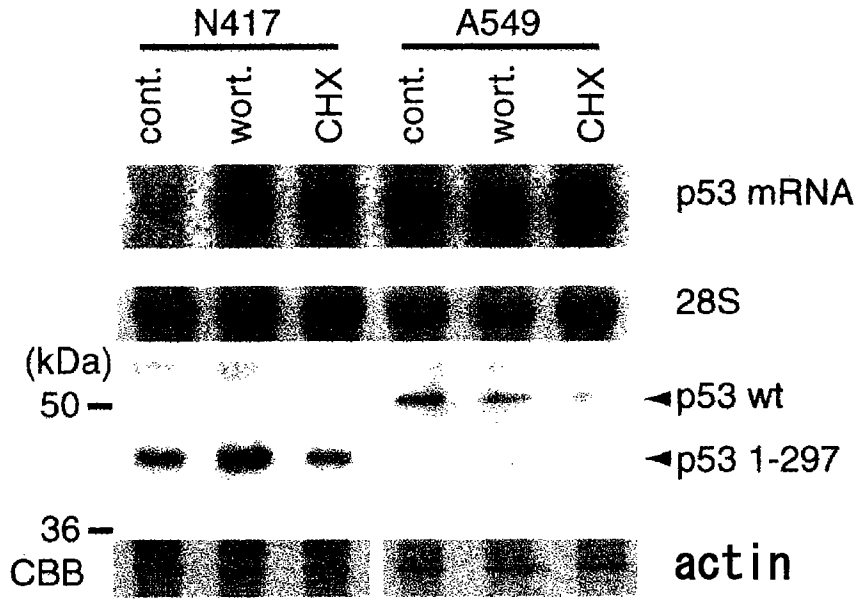
[Figure 29]



[Figure 30]



[Figure 31]



[Figure 32]